



STIC SEARCH RESULTS

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or contact:

Mary Hale, Information Branch Supervisor
308-4258, CM1-1E01

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library/CM1 - Circ. Desk



THIS PAGE BLANK (USPTO)

PT vaccae have enhanced activity in the presence of organic solvents and
 PT are useful for producing alcohols from ketone raw materials

PS Example 1; Page 23-24; 42pp; English.

CC This invention relates to a mutant form of the Mycobacterium vaccae
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents the Mycobacterium vaccae
 CC formate dehydrogenase protein of the invention. This sequence was used
 CC to generate the mutant proteins of the invention.

XX Sequence 401 AA;

SO Query Match 100.0%; Score 2135; DB 23; Length 401;
 Best Local Similarity 100.0%; Pred. No. 1.5e-212;
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARVLCVLYDDPVDGPKYKTAARDLPRKIDHYPGGQILPTPKAIDFTPGQLGSGSELGL 60
 DB 1 MARVLCVLYDDPVDGPKYKTAARDLPRKIDHYPGGQILPTPKAIDFTPGQLGSGSELGL 60
 QY RPYLESNGHTLVVTSKDGPSVFERELVDADVISOFPFPAVLTPERIAKANKLKALT 120
 DB RPYLESNGHTLVVTSKDGPSVFERELVDADVISOFPFPAVLTPERIAKANKLKALT 120
 QY 121 AGGSDHVDQSAIDRNVYTAETTCNSISVAEHVMMILSVRNTLPSEHMARKGWN 180
 DB 121 AGGSDHVDQSAIDRNVYTAETTCNSISVAEHVMMILSVRNTLPSEHMARKGWN 180
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDHRRLPESYEKELNTLWHA 240
 DB 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDHRRLPESYEKELNTLWHA 240
 QY 241 TREDMYPCDVVTLNCPLEHETHEMINDETLKFRCGAYIVNTARGKLCDRDVARALES 300
 DB 241 TREDMYPCDVVTLNCPLEHETHEMINDETLKFRCGAYIVNTARGKLCDRDVARALES 300
 QY 301 GRLAGYAGDVWFPQAPKDPHPRMTMPYNGTPIHSGTTLTAQARYAAGTREILECFEGR 360
 DB 301 GRLAGYAGDVWFPQAPKDPHPRMTMPYNGTPIHSGTTLTAQARYAAGTREILECFEGR 360
 QY 361 PIRDEYLIYOGGALAGTAGHASYSKGNATGSEEAARFKKAV 401
 DB 361 PIRDEYLIYOGGALAGTAGHASYSKGNATGSEEAARFKKAV 401

RESULT 2

AAU99112 standard; Protein; 401 AA.

AAU99112;

24-SEP-2002 (first entry)

DE Mycobacterium vaccae formate dehydrogenase C146S mutant.

KW Formate dehydrogenase; alcohol production; ketone;
 mutant; mutlein.

OS Mycobacterium vaccae.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 146

XX /note= "Wild type Cys substituted by Ser"

EP1211316-A1.

XX

PD 05-JUN-2002.

XX 27-NOV-2001; 2001EP-0128170.

XX 29-NOV-2000; 2000JP-0363894.

PR 24-AUG-2001; 2001JP-0254631.

PA (DAI) DAICEL CHEM IND LTD.

PI Mitsuhashi K, Yamamoto H, Kimoto N;

DR WPI; 2002-464925/50.

PT New mutant forms of formate dehydrogenase derived from Mycobacterium

PT vaccae have enhanced activity in the presence of organic solvents and

PT are useful for producing alcohols from ketone raw materials

PS Claim 8; Page -; 42pp; English.

CC This invention relates to a mutant form of the Mycobacterium vaccae
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a Mycobacterium vaccae
 CC formate dehydrogenase protein mutant of the invention.

XX Sequence 401 AA;

SO Query Match 99.5%; Score 2125; DB 23; Length 401;
 Best Local Similarity 99.8%; Pred. No. 1.6e-211;
 Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARVLCVLYDDPVDGPKYKTAARDLPRKIDHYPGGQILPTPKAIDFTPGQLGSGSELGL 60
 DB 1 MARVLCVLYDDPVDGPKYKTAARDLPRKIDHYPGGQILPTPKAIDFTPGQLGSGSELGL 60
 QY RPYLESNGHTLVVTSKDGPSVFERELVDADVISOFPFPAVLTPERIAKANKLKALT 120
 DB RPYLESNGHTLVVTSKDGPSVFERELVDADVISOFPFPAVLTPERIAKANKLKALT 120
 QY 121 AGGSDHVDQSAIDRNVYTAETTCNSISVAEHVMMILSVRNTLPSEHMARKGWN 180
 DB 121 AGGSDHVDQSAIDRNVYTAETTCNSISVAEHVMMILSVRNTLPSEHMARKGWN 180
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDHRRLPESYEKELNTLWHA 240
 DB 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDHRRLPESYEKELNTLWHA 240
 QY 241 TREDMYPCDVVTLNCPLEHETHEMINDETLKFRCGAYIVNTARGKLCDRDVARALES 300
 DB 241 TREDMYPCDVVTLNCPLEHETHEMINDETLKFRCGAYIVNTARGKLCDRDVARALES 300
 QY 301 GRLAGYAGDVWFPQAPKDPHPRMTMPYNGTPIHSGTTLTAQARYAAGTREILECFEGR 360
 DB 301 GRLAGYAGDVWFPQAPKDPHPRMTMPYNGTPIHSGTTLTAQARYAAGTREILECFEGR 360
 QY 361 PIRDEYLIYOGGALAGTAGHASYSKGNATGSEEAARFKKAV 401
 DB 361 PIRDEYLIYOGGALAGTAGHASYSKGNATGSEEAARFKKAV 401

RESULT 3

AAU99113 standard; Protein; 401 AA.

AAU99113;

24-SEP-2002 (first entry)

DE Mycobacterium vaccae formate dehydrogenase C256S mutant.

XX Formate dehydrogenase; alcohol production; ketone;
 KW mutant; muteln.
 XX
 XX Mycobacterium vaccae.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 256
 FT /note= "Wild type Cys substituted by Ser"
 XX
 XX EPI211316-A1.
 XX
 PD 05-JUN-2002.
 XX
 PF 27-NOV-2001; 2001EP-0128170.
 XX
 PR 29-NOV-2000; 2000JP-0363894.
 PR 24-AUG-2001; 2001JP-0254631.
 XX
 PA (DAIICEL CHEM IND LTD.
 XX
 PI Mitsuhashi K, Yamamoto H, Kimoto N;
 XX
 DR WPI; 2002-464925/50.
 XX
 PT New mutant forms of formate dehydrogenase derived from Mycobacterium
 PT vaccae have enhanced activity in the presence of organic solvents and
 PT are useful for producing alcohols from ketone raw materials
 XX
 XX Claim 8; Page -: 42pp; English.
 XX
 PS This invention relates to a mutant form of the Mycobacterium vaccae
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a Mycobacterium vaccae
 CC formate dehydrogenase protein mutant of the invention.
 CC
 SQ Sequence 401 AA:
 Query Match 99.5%; Score 2125; DB 23; Length 401;
 Best Local Similarity 99.8%; Pred. No. 1.6e-211;
 Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MAKVLCVLYDDPVDGPKTYARDLPRKIDHYPGQILPTPKAIDFTPGQLGSVSGELGL 60
 DB 1 MAKVLCVLYDDPVDGPKTYARDLPRKIDHYPGQILPTPKAIDFTPGQLGSVSGELGL 60
 QY 61 RPYLESNGHTLVVTSDDKGPDSVFEREVLVDADVVISQFPWPAYLTPERIAKKNKLALT 120
 DB 61 RPYLESNGHTLVVTSDDKGPDSVFEREVLVDADVVISQFPWPAYLTPERIAKKNKLALT 120
 QY 121 AGISGDHVDLOSADIRNTVAEVTYCNISVAEHVMMILSVRYNYLSHEHARKGNMI 180
 DB 121 AGISGDHVDLOSADIRNTVAEVTYCNISVAEHVMMILSVRYNYLSHEHARKGNMI 180
 QY 181 ADVSHAYDLEAMHVGTAAGRIGLAVLRRLAPFDVHLHYTRHRLPESVEKELMLTWH 240
 DB 181 ADVSHAYDLEAMHVGTAAGRIGLAVLRRLAPFDVHLHYTRHRLPESVEKELMLTWH 240
 QY 241 TREDMYRVCDDVYTLNCPILHPETEHMIINDETLLKLFKRGAYIVNTARGKLCDRDAVARALES 300
 DB 241 TREDMYRVCDDVYTLNCPILHPETEHMIINDETLLKLFKRGAYIVNTARGKLCDRDAVARALES 300
 QY 301 GRIGYAGDVYFPPAPKDHPRMTPYNGMTPHISGTLTQAARYAAGTRRELECFEGR 360
 DB 301 GRIGYAGDVYFPPAPKDHPRMTPYNGMTPHISGTLTQAARYAAGTRRELECFEGR 360
 QY 361 PIRDEYLIYOGGALAGTGAHYSKGNATGSGSEAKFKKAV 401

DB 361 PIRDEYLIYOGGALAGTGAHYSKGNATGSGSEAKFKKAV 401
 RESULT 4
 ID AAU99115 standard; Protein; 401 AA.
 XX
 XX AAU99115;
 XX
 DT 24-SEP-2002 (first entry)
 XX
 DE Mycobacterium vaccae formate dehydrogenase C256v mutant.
 XX
 DE Formate dehydrogenase; alcohol production; ketone;
 KW mutant; muteln.
 XX
 OS Mycobacterium vaccae.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 256
 FT /note= "Wild type Cys substituted by Val"
 XX
 XX EPI211316-A1.
 XX
 PD 05-JUN-2002.
 XX
 PF 27-NOV-2001; 2001EP-0128170.
 XX
 PR 29-NOV-2000; 2000JP-0363894.
 PR 24-AUG-2001; 2001JP-0254631.
 XX
 PA (DAIICEL CHEM IND LTD.
 XX
 PI Mitsuhashi K, Yamamoto H, Kimoto N;
 XX
 DR WPI; 2002-464925/50.
 XX
 PT New mutant forms of formate dehydrogenase derived from Mycobacterium
 PT vaccae have enhanced activity in the presence of organic solvents and
 PT are useful for producing alcohols from ketone raw materials
 XX
 XX Claim 8; Page -: 42pp; English.
 XX
 PS This invention relates to a mutant form of the Mycobacterium vaccae
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a Mycobacterium vaccae
 CC formate dehydrogenase protein mutant of the invention.
 CC
 SQ Sequence 401 AA:
 Query Match 99.5%; Score 2125; DB 23; Length 401;
 Best Local Similarity 99.8%; Pred. No. 1.6e-211;
 Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MAKVLCVLYDDPVDGPKTYARDLPRKIDHYPGQILPTPKAIDFTPGQLGSVSGELGL 60
 DB 1 MAKVLCVLYDDPVDGPKTYARDLPRKIDHYPGQILPTPKAIDFTPGQLGSVSGELGL 60
 QY 61 RPYLESNGHTLVVTSDDKGPDSVFEREVLVDADVVISQFPWPAYLTPERIAKKNKLALT 120
 DB 61 RPYLESNGHTLVVTSDDKGPDSVFEREVLVDADVVISQFPWPAYLTPERIAKKNKLALT 120
 QY 121 AGISGDHVDLOSADIRNTVAEVTYCNISVAEHVMMILSVRYNYLSHEHARKGNMI 180
 DB 121 AGISGDHVDLOSADIRNTVAEVTYCNISVAEHVMMILSVRYNYLSHEHARKGNMI 180

PS Claim 8; Page -: 42pp; English.
XX This invention relates to a mutant form of the Mycobacterium vaccae
CC formate dehydrogenase protein which has strong formate dehydrogenase
CC activity in the presence of an organic solvent. The mutant formate
CC dehydrogenase protein of the invention may be used to produce alcohols
CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
CC proteins, the mutant polypeptides of the invention retain high enzymatic
CC activity in the presence of the alcohol product and so provides a higher
CC product yield. The present sequence represents a Mycobacterium vaccae
CC formate dehydrogenase protein mutant of the invention.
XX
SQ Sequence 401 AA;
Query Match 99.1%; Score 2116; DB 23; Length 401;
Best Local Similarity 99.5%; Pred. No. 1.4e-210;
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAKVLCVLYDDPVDVGYKRTYARDDLPRKIDHYPGQIILPTPKAIDFTPGQLGVSGETGL 60
DB 1 MAKVLCVLYDDPVDVGYKRTYARDDLPRKIDHYPGQIILPTPKAIDFTPGQLGVSGETGL 60
QY 61 RPYLESNGHTLVYTSDDKGPDSVFERELVDADVVISQPFMPAYLTPERIAKAKMLKALT 120
DB 61 RPYLESNGHTLVYTSDDKGPDSVFERELVDADVVISQPFMPAYLTPERIAKAKMLKALT 120
QY 121 AGISDHDVLDOSADIRNVTAEVYTCNSISVAEHVMMIISLVNRYLPSHEMARKGWN 180
DB 121 AGISDHDVLDOSADIRNVTAEVYTCNSISVAEHVMMIISLVNRYLPSHEMARKGWN 180
QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPDPVHLHYDRHRLPESVEKEINLTWHA 240
DB 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPDPVHLHYDRHRLPESVEKEINLTWHA 240
QY 241 TREMYPCDVVYTNCLPHETEMINDETLKFKRGAYIYNTARGKICDDAVARALES 300
DB 241 TREMYPCDVVYTNCLPHETEMINDETLKFKRGAYIYNTARGKICDDAVARALES 300
QY 301 GRLAGYAGDVWFPPAPKDHPRMTPYNGMTPHISGTTLTQAARYAAGTRILECFEGR 360
DB 301 GRLAGYAGDVWFPPAPKDHPRMTPYNGMTPHISGTTLTQAARYAAGTRILECFEGR 360
QY 361 PIRDEYLIYOGALAGTGAHSHYSKGNATGSEEAKEKRAV 401
DB 361 PIRDEYLIYOGALAGTGAHSHYSKGNATGSEEAKEKRAV 401
RESULT 7
ID AAU99117 standard; Protein: 401 AA.
AC AAU99117;
XX 24-SEP-2002 (first entry)
DE Mycobacterium vaccae formate dehydrogenase C6A/C256V mutant.
XX
KW Formate dehydrogenase; alcohol production; ketone;
XX mutant; mutcin.
OS Mycobacterium vaccae.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 6 /note- "Wild type Cys substituted by Ala"
FT Misc-difference 256 /note- "Wild type Cys substituted by Val"
FT
XX EPI211316-A1.
XX
XX 05-JUN-2002.
XX

PF 27-NOV-2001; 2001EP-0128170.
XX
PR 29-NOV-2000; 2000JP-0363894.
PR 24-AUG-2001; 2001JP-0254631.
XX
PA (DAI) DAICEL CHEM IND LTD.
PI Mitsuhashi K, Yamamoto H, Kimoto N;
XX WPI: 2002-464925/50.
XX
PT New mutant forms of formate dehydrogenase derived from Mycobacterium
PT vaccae have enhanced activity in the presence of organic solvents and
PT are useful for producing alcohols from ketone raw materials
XX
PS Claim 8; Page -: 42pp; English.
XX
XX This invention relates to a mutant form of the Mycobacterium vaccae
CC formate dehydrogenase protein which has strong formate dehydrogenase
CC activity in the presence of an organic solvent. The mutant formate
CC dehydrogenase protein of the invention may be used to produce alcohols
CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
CC proteins, the mutant polypeptides of the invention retain high enzymatic
CC activity in the presence of the alcohol product and so provides a higher
CC product yield. The present sequence represents a Mycobacterium vaccae
CC formate dehydrogenase protein mutant of the invention.
XX
SQ Sequence 401 AA;
Query Match 99.1%; Score 2116; DB 23; Length 401;
Best Local Similarity 99.5%; Pred. No. 1.4e-210;
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAKVLCVLYDDPVDVGYKRTYARDDLPRKIDHYPGQIILPTPKAIDFTPGQLGVSGETGL 60
DB 1 MAKVLCVLYDDPVDVGYKRTYARDDLPRKIDHYPGQIILPTPKAIDFTPGQLGVSGETGL 60
QY 61 RPYLESNGHTLVYTSDDKGPDSVFERELVDADVVISQPFMPAYLTPERIAKAKMLKALT 120
DB 61 RPYLESNGHTLVYTSDDKGPDSVFERELVDADVVISQPFMPAYLTPERIAKAKMLKALT 120
QY 121 AGISDHDVLDOSADIRNVTAEVYTCNSISVAEHVMMIISLVNRYLPSHEMARKGWN 180
DB 121 AGISDHDVLDOSADIRNVTAEVYTCNSISVAEHVMMIISLVNRYLPSHEMARKGWN 180
QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPDPVHLHYDRHRLPESVEKEINLTWHA 240
DB 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPDPVHLHYDRHRLPESVEKEINLTWHA 240
QY 241 TREMYPCDVVYTNCLPHETEMINDETLKFKRGAYIYNTARGKICDDAVARALES 300
DB 241 TREMYPCDVVYTNCLPHETEMINDETLKFKRGAYIYNTARGKICDDAVARALES 300
QY 301 GRLAGYAGDVWFPPAPKDHPRMTPYNGMTPHISGTTLTQAARYAAGTRILECFEGR 360
DB 301 GRLAGYAGDVWFPPAPKDHPRMTPYNGMTPHISGTTLTQAARYAAGTRILECFEGR 360
QY 361 PIRDEYLIYOGALAGTGAHSHYSKGNATGSEEAKEKRAV 401
DB 361 PIRDEYLIYOGALAGTGAHSHYSKGNATGSEEAKEKRAV 401
RESULT 8
ID AAU99109 standard; Protein: 401 AA.
AC AAU99109;
XX 24-SEP-2002 (first entry)
DE Mycobacterium vaccae formate dehydrogenase C6V/C256S mutant.
XX
KW Formate dehydrogenase; alcohol production; ketone;
XX

```

KM mutant; mutein.
XX
OS Mycobacterium vaccae.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 6 /note= "Wild type Cys substituted by Val"
FT 256 /note= "Wild type Cys substituted by Ser"
FT Misc-difference 256 /note= "Wild type Cys substituted by Ser"
XX
XX EP1211316-A1.
XX
XX 05-JUN-2002.
XX
XX 27-NOV-2001, 2001EP-0128170.
XX
XX 29-NOV-2000, 2000JP-0363894.
XX 24-AUG-2001, 2001JP-0254631.
XX
XX (DAIL ) DAICEL CHEM IND LTD.
XX
XX Mitsuhashi K, Yamamoto H, Kimoto N;
XX
XX WPI; 2002-464925/50.
XX
XX
XX New mutant forms of formate dehydrogenase derived from Mycobacterium
XX PT vaccae have enhanced activity in the presence of organic solvents and
XX PT are useful for producing alcohols from ketone raw materials
XX
XX Claim 8; Page -: 42pp; English.
XX
XX This invention relates to a mutant form of the Mycobacterium vaccae
XX CC formate dehydrogenase protein which has strong formate dehydrogenase
XX CC activity in the presence of an organic solvent. The mutant formate
XX CC dehydrogenase protein of the invention may be used to produce alcohols
XX CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
XX CC proteins, the mutant polypeptides of the invention retain high enzymatic
XX CC activity in the presence of the alcohol product and so provides a higher
XX CC product yield. The present sequence represents a Mycobacterium vaccae
XX CC formate dehydrogenase protein mutant of the invention.
XX
XX Sequence 401 AA:
XX
Query Match 99.1%; Score 2115; DB 23; Length 401;
Best Local Similarity 99.5%; Pred. No. 1.7e-210;
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAKVLCVLYDDPVDGYPKTYARDLPKIDHYRGQILPFPKAIIDFTPGQLLSVSGELGL 60
DB 1 MAKVLYLVYDDPVDGYPKTYARDLPKIDHYRGQILPFPKAIIDFTPGQLLSVSGELGL 60
QY RPYLESNGHTLVVTSKDGPDVSFERELVDADVVISQFPWPAVLTPERIAKKAKNKLAL 120
DB RPYLESNGHTLVVTSKDGPDVSFERELVDADVVISQFPWPAVLTPERIAKKAKNKLAL 120
QY 61 RPYLESNGHTLVVTSKDGPDVSFERELVDADVVISQFPWPAVLTPERIAKKAKNKLAL 120
DB 61 RPYLESNGHTLVVTSKDGPDVSFERELVDADVVISQFPWPAVLTPERIAKKAKNKLAL 120
QY 121 AGISGDHVDLSAIDBNVTVAEVTYCNSSISVAEHVVMILSLVRNLPSEHMAKKGWNI 180
DB 121 AGISGDHVDLSAIDBNVTVAEVTYCNSSISVAEHVVMILSLVRNLPSEHMAKKGWNI 180
QY 181 ADCVSNAYDLEAHNVGTVAAGRTGLAVLRRLAFEDVHLHTDHRRLPESYEKELNTWRA 240
DB 181 ADCVSNAYDLEAHNVGTVAAGRTGLAVLRRLAFEDVHLHTDHRRLPESYEKELNTWRA 240
QY 241 TREDMTPVCDVYTLNPLRPETEHMINDETLPKFKRGATVYNTARSKLCDRDVARALSS 300
DB 241 TREDMTPVCDVYTLNPLRPETEHMINDETLPKFKRGATVYNTARSKLCDRDVARALSS 300
QY 301 GRLAGYAGDVWFPQAPKDPWRTMPYNGMTPHISGTTTLAQAARYAAGTRELIECFEGR 360
DB 301 GRLAGYAGDVWFPQAPKDPWRTMPYNGMTPHISGTTTLAQAARYAAGTRELIECFEGR 360
QY 361 PIRDEYLIYOGGALAGTGAHSYKGNATGSEBAAFKRAV 401

```

```

DB 361 PIRDEYLIYOGGALAGTGAHSYKGNATGSEBAAFKRAV 401
|||||
RESULT 9
AA099111
ID AA099111 standard; Protein; 401 AA.
XX
AC AA099111;
XX
DT 24-SEP-2002 (first entry)
XX
XX Mycobacterium vaccae formate dehydrogenase C6S/C256V mutant.
XX
XX Formate dehydrogenase; alcohol production; ketone;
XX mutant; mutein.
XX
XX Mycobacterium vaccae.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 6 /note= "Wild type Cys substituted by Ser"
FT 256 /note= "Wild type Cys substituted by Val"
FT Misc-difference 256 /note= "Wild type Cys substituted by Val"
XX
XX EP1211316-A1.
XX
XX 05-JUN-2002.
XX
XX 27-NOV-2001, 2001EP-0128170.
XX
XX 29-NOV-2000, 2000JP-0363894.
XX 24-AUG-2001, 2001JP-0254631.
XX
XX (DAIL ) DAICEL CHEM IND LTD.
XX
XX Mitsuhashi K, Yamamoto H, Kimoto N;
XX
XX WPI; 2002-464925/50.
XX
XX
XX New mutant forms of formate dehydrogenase derived from Mycobacterium
XX PT vaccae have enhanced activity in the presence of organic solvents and
XX PT are useful for producing alcohols from ketone raw materials
XX
XX Claim 8; Page -: 42pp; English.
XX
XX This invention relates to a mutant form of the Mycobacterium vaccae
XX CC formate dehydrogenase protein which has strong formate dehydrogenase
XX CC activity in the presence of an organic solvent. The mutant formate
XX CC dehydrogenase protein of the invention may be used to produce alcohols
XX CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
XX CC proteins, the mutant polypeptides of the invention retain high enzymatic
XX CC activity in the presence of the alcohol product and so provides a higher
XX CC product yield. The present sequence represents a Mycobacterium vaccae
XX CC formate dehydrogenase protein mutant of the invention.
XX
XX Sequence 401 AA:
XX
Query Match 99.1%; Score 2115; DB 23; Length 401;
Best Local Similarity 99.5%; Pred. No. 1.7e-210;
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAKVLCVLYDDPVDGYPKTYARDLPKIDHYRGQILPFPKAIIDFTPGQLLSVSGELGL 60
DB 1 MAKVLYLVYDDPVDGYPKTYARDLPKIDHYRGQILPFPKAIIDFTPGQLLSVSGELGL 60
QY 61 RPYLESNGHTLVVTSKDGPDVSFERELVDADVVISQFPWPAVLTPERIAKKAKNKLAL 120
DB 61 RPYLESNGHTLVVTSKDGPDVSFERELVDADVVISQFPWPAVLTPERIAKKAKNKLAL 120
QY 121 AGISGDHVDLSAIDBNVTVAEVTYCNSSISVAEHVVMILSLVRNLPSEHMAKKGWNI 180
DB 121 AGISGDHVDLSAIDBNVTVAEVTYCNSSISVAEHVVMILSLVRNLPSEHMAKKGWNI 180

```


Db 121 AGIGSDHVDLSAIDRNVTAEVYTCNSISVAEHVMMILSLVRNYLPSEHEMARGGWN 180
Qy 181 ADCVSHAYDLEAMHGVTAAGRIGLAVLRRLAPFDVHLHYTDNRRLPESVEKELNLTWHA 240
Db 181 ADCVSHAYDLEAMHGVTAAGRIGLAVLRRLAPFDVHLHYTDNRRLPESVEKELNLTWHA 240
Qy 241 TREDMYPVCADVTLNCPILHPETEHMINDETLLKFRGAYIYNTARGKLCDDDAVARALES 300
Db 241 TREDMYPVCADVTLNCPILHPETEHMINDETLLKFRGAYIYNTARGKLCDDDAVARALES 300
Qy 301 GRLAGYAGDVWFPPAPKDPHWRTPYNGMTPHISGTTLTQAARYAACTREILCECFEGR 360
Db 301 GRLAGYAGDVWFPPAPKDPHWRTPYNGMTPHISGTTLTQAARYAACTREILCECFEGR 360
Qy 361 PIRDEYLLVGGALAGTGAHSYSGNATGSGSEAAKFKKAV 401
Db 361 PIRDEYLLVGGALAGTGAHSYSGNATGSGSEAAKFKKAV 401

RESULT 10

AAU99114

ID AAU99114 standard; Protein: 401 AA.

AC AAU99114:

DT 24-SEP-2002 (first entry)

De Mycobacterium vaccae formate dehydrogenase C146S/C256S mutant.

Kw Formate dehydrogenase; alcohol production; ketone;

Kw mutant; mutein.

OS Mycobacterium vaccae.

OS Synthetic.

Key Location/Qualifiers

FH Misc-difference 146

FT /note= "Wild type Cys substituted by Ser"

FT Misc-difference 256

FT /note= "Wild type Cys substituted by Ser"

PN EP1211316-A1.

PD 05-JUN-2002.

PF 27-NOV-2001; 2001EP-0128170.

PR 29-NOV-2000; 2000JP-0363894.

PR 24-AUG-2001; 2001JP-0254631.

PA (DAIL) DAICEL CHEM IND LTD.

PI Mitsuhashi K, Yamamoto H, Kimoto N;

DR WPI; 2002-464925/50.

PT New mutant forms of formate dehydrogenase derived from Mycobacterium
PT vaccae have enhanced activity in the presence of organic solvents and
PT are useful for producing alcohols from ketone raw materials

PS Claim 8; Page -: 42pp; English.

CC This invention relates to a mutant form of the Mycobacterium vaccae
CC formate dehydrogenase protein which has strong formate dehydrogenase
CC activity in the presence of an organic solvent. The mutant formate
CC dehydrogenase protein of the invention may be used to produce alcohols
CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
CC proteins, the mutant polypeptides of the invention retain high enzymatic
CC activity in the presence of the alcohol product and so provides a higher
CC product yield. The present sequence represents a Mycobacterium vaccae
CC formate dehydrogenase protein mutant of the invention.

Sequence 401 AA;

Query Match 99.1%; Score 2115; DB 23; Length 401;
Best Local Similarity 99.5%; Pred. No. 1.7e-210;
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAKVLCVLYDDPVGVGPKTYARDDLPRKIDHPGGLTTPRAIDFTPGQLGVSVELGL 60
Db 1 MAKVLCVLYDDPVGVGPKTYARDDLPRKIDHPGGLTTPRAIDFTPGQLGVSVELGL 60
Qy 61 RPYLESNGHTLVTSDDKGPDSVEERELVDADVVISQFPWPAYLTPERIAKAKNLKALT 120
Db 61 RPYLESNGHTLVTSDDKGPDSVEERELVDADVVISQFPWPAYLTPERIAKAKNLKALT 120
Qy 121 AGIGSDHVDLSAIDRNVTAEVYTCNSISVAEHVMMILSLVRNYLPSEHEMARGGWN 180
Db 121 AGIGSDHVDLSAIDRNVTAEVYTCNSISVAEHVMMILSLVRNYLPSEHEMARGGWN 180
Qy 181 ADCVSHAYDLEAMHGVTAAGRIGLAVLRRLAPFDVHLHYTDNRRLPESVEKELNLTWHA 240
Db 181 ADCVSHAYDLEAMHGVTAAGRIGLAVLRRLAPFDVHLHYTDNRRLPESVEKELNLTWHA 240
Qy 241 TREDMYPVCADVTLNCPILHPETEHMINDETLLKFRGAYIYNTARGKLCDDDAVARALES 300
Db 241 TREDMYPVCADVTLNCPILHPETEHMINDETLLKFRGAYIYNTARGKLCDDDAVARALES 300
Qy 301 GRLAGYAGDVWFPPAPKDPHWRTPYNGMTPHISGTTLTQAARYAACTREILCECFEGR 360
Db 301 GRLAGYAGDVWFPPAPKDPHWRTPYNGMTPHISGTTLTQAARYAACTREILCECFEGR 360
Qy 361 PIRDEYLLVGGALAGTGAHSYSGNATGSGSEAAKFKKAV 401
Db 361 PIRDEYLLVGGALAGTGAHSYSGNATGSGSEAAKFKKAV 401

RESULT 11

AAU99116

ID AAU99116 standard; Protein: 401 AA.

AC AAU99116:

DT 24-SEP-2002 (first entry)

De Mycobacterium vaccae formate dehydrogenase C146S/C256V mutant.

Kw Formate dehydrogenase; alcohol production; ketone;

Kw mutant; mutein.

OS Mycobacterium vaccae.

OS Synthetic.

Key Location/Qualifiers

FH Misc-difference 146

FT /note= "Wild type Cys substituted by Ser"

FT Misc-difference 256

FT /note= "Wild type Cys substituted by Val"

PN EP1211316-A1.

PD 05-JUN-2002.

PF 27-NOV-2001; 2001EP-0128170.

PR 29-NOV-2000; 2000JP-0363894.

PR 24-AUG-2001; 2001JP-0254631.

PA (DAIL) DAICEL CHEM IND LTD.

PI Mitsuhashi K, Yamamoto H, Kimoto N;

DR WPI; 2002-464925/50.

PT New mutant forms of formate dehydrogenase derived from Mycobacterium
PT vaccae have enhanced activity in the presence of organic solvents and

PT are useful for producing alcohols from ketone raw materials -
 XX Claim 8; Page -: 42pp; English.
 XX
 CC This invention relates to a mutant form of the *Mycobacterium vaccae*
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a *Mycobacterium vaccae*
 CC formate dehydrogenase protein mutant of the invention.
 CC
 XX
 SQ Sequence 401 AA;
 Query Match 99.1%; Score 2115; DB 23; Length 401;
 Best Local Similarity 99.5%; Pred. No. 1.7e-210;
 Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MAKVLCVLYDDPYDGYPKTYARDLPRKIDHYPGQILPPPKAIDFTPGQLGSGVSGELG 60
 DB 1 MAKVLCVLYDDPYDGYPKTYARDLPRKIDHYPGQILPPPKAIDFTPGQLGSGVSGELG 60
 QY 61 RPLLESNGHTLVYTSKDGSDSVFERELVDADYVISOPFPAYLTPERIKAKNKLALT 120
 DB 61 RPLLESNGHTLVYTSKDGSDSVFERELVDADYVISOPFPAYLTPERIKAKNKLALT 120
 QY 121 AGIGSDHVDLSAIDRNVTVAEVTYCNISVAEHVMMILSLVRNLYPSHEMARKGWNI 180
 DB 121 AGIGSDHVDLSAIDRNVTVAEVTYCNISVAEHVMMILSLVRNLYPSHEMARKGWNI 180
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDRHRLPESVEKELNLTWHA 240
 DB 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDRHRLPESVEKELNLTWHA 240
 QY 241 TREDMPVPCDVYTLNPLHPETEHEMIDETLKFRCRAYIVNTARGKLCRDVAARLES 300
 DB 241 TREDMPVPCDVYTLNPLHPETEHEMIDETLKFRCRAYIVNTARGKLCRDVAARLES 300
 QY 301 GRLAGYAGDVWFPQAPKDHPRMTMPYNGMTPHISGTTLTAQARYAAGTREILECFEGR 360
 DB 301 GRLAGYAGDVWFPQAPKDHPRMTMPYNGMTPHISGTTLTAQARYAAGTREILECFEGR 360
 QY 361 PIRDEVLYVOGALAGTAGHASYSGKNGATGSEEAKEKAV 401
 DB 361 PIRDEVLYVOGALAGTAGHASYSGKNGATGSEEAKEKAV 401
 RESULT 12
 AAU99119 standard; Protein: 401 AA.
 XX
 AC AAU99119;
 XX
 DT 24-SEP-2002 (first entry)
 XX
 DE *Mycobacterium vaccae* formate dehydrogenase C6a/C146a/C256v mutant.
 XX
 KM Formate dehydrogenase; alcohol production; ketone;
 KM mutant; mutlein.
 OS *Mycobacterium vaccae*.
 OS synthetic.
 XX
 FH Key Location/Qualifiers
 FT MISC-difference 6 /note- "Wild type Cys substituted by Ala"
 FT MISC-difference 146 /note- "Wild type Cys substituted by Ala"
 FT MISC-difference 256 /note- "Wild type Cys substituted by Val"
 XX

PN EP1211316-A1.
 XX
 PD 05-JUN-2002.
 XX
 PF 27-NOV-2001; 2001EP-0128170.
 XX
 PR 29-NOV-2000; 2000JP-0363894.
 PR 24-AUG-2001; 2001JP-0254631.
 XX
 PA (DAIIC) DAICEL CHEM IND LTD.
 XX
 PI Mitsuhashi K, Yamamoto H, Kimoto N;
 XX
 DR WPI; 2002-464925/50.
 XX
 PT New mutant forms of formate dehydrogenase derived from *Mycobacterium*
 PT *vaccae* have enhanced activity in the presence of organic solvents and
 PT are useful for producing alcohols from ketone raw materials -
 XX
 XX Claim 8; Page -: 42pp; English.
 XX
 CC This invention relates to a mutant form of the *Mycobacterium vaccae*
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a *Mycobacterium vaccae*
 CC formate dehydrogenase protein mutant of the invention.
 CC
 XX
 SQ Sequence 401 AA;
 Query Match 98.7%; Score 2107; DB 23; Length 401;
 Best Local Similarity 99.3%; Pred. No. 1.2e-209;
 Matches 398; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MAKVLCVLYDDPYDGYPKTYARDLPRKIDHYPGQILPPPKAIDFTPGQLGSGVSGELG 60
 DB 1 MAKVLCVLYDDPYDGYPKTYARDLPRKIDHYPGQILPPPKAIDFTPGQLGSGVSGELG 60
 QY 61 RPLLESNGHTLVYTSKDGSDSVFERELVDADYVISOPFPAYLTPERIKAKNKLALT 120
 DB 61 RPLLESNGHTLVYTSKDGSDSVFERELVDADYVISOPFPAYLTPERIKAKNKLALT 120
 QY 121 AGIGSDHVDLSAIDRNVTVAEVTYCNISVAEHVMMILSLVRNLYPSHEMARKGWNI 180
 DB 121 AGIGSDHVDLSAIDRNVTVAEVTYCNISVAEHVMMILSLVRNLYPSHEMARKGWNI 180
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDRHRLPESVEKELNLTWHA 240
 DB 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDRHRLPESVEKELNLTWHA 240
 QY 241 TREDMPVPCDVYTLNPLHPETEHEMIDETLKFRCRAYIVNTARGKLCRDVAARLES 300
 DB 241 TREDMPVPCDVYTLNPLHPETEHEMIDETLKFRCRAYIVNTARGKLCRDVAARLES 300
 QY 301 GRLAGYAGDVWFPQAPKDHPRMTMPYNGMTPHISGTTLTAQARYAAGTREILECFEGR 360
 DB 301 GRLAGYAGDVWFPQAPKDHPRMTMPYNGMTPHISGTTLTAQARYAAGTREILECFEGR 360
 QY 361 PIRDEVLYVOGALAGTAGHASYSGKNGATGSEEAKEKAV 401
 DB 361 PIRDEVLYVOGALAGTAGHASYSGKNGATGSEEAKEKAV 401
 RESULT 13
 AAU99118 standard; Protein: 401 AA.
 XX
 AC AAU99118;
 XX
 DT 24-SEP-2002 (first entry)

XX Mycobacterium vaccae formate dehydrogenase C6A/C146S/C256V mutant.
DE
XX Formate dehydrogenase; alcohol production; ketone;
KW mutant; mutlein.
XX
OS Mycobacterium vaccae.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 6 /note= "Wild type Cys substituted by Ala"
FT Misc-difference 146 /note= "Wild type Cys substituted by Ser"
FT Misc-difference 256 /note= "Wild type Cys substituted by Val"
FT
XX
XX EPI211316-A1.
XX
XX 05-JUN-2002.
XX
XX 27-NOV-2001; 2001EP-0128170.
XX
XX 29-NOV-2000; 2000JP-0363894.
XX 24-AUG-2001; 2001JP-0254631.
XX
XX (DAIIL) DAICEL CHEM IND LTD.
XX
XX Mitsuhashi K, Yamamoto H, Kimoto N;
XX
XX WPI; 2002-464925/50.
XX
XX
XX New mutant forms of formate dehydrogenase derived from Mycobacterium
PT vaccae have enhanced activity in the presence of organic solvents and
PT are useful for producing alcohols from ketone raw materials
XX
XX
XX Claim 8; Page -: 42pp; English.
XX
XX This invention relates to a mutant form of the Mycobacterium vaccae
CC formate dehydrogenase protein which has strong formate dehydrogenase
CC activity in the presence of an organic solvent. The mutant formate
CC dehydrogenase protein of the invention may be used to produce alcohols
CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
CC proteins, the mutant polypeptides of the invention retain high enzymatic
CC activity in the presence of the alcohol product and so provides a higher
CC product yield. The present sequence represents a Mycobacterium vaccae
CC formate dehydrogenase protein mutant of the invention.
CC
XX
SQ Sequence 401 AA;
Query Match 98.6%; Score 2106; DB 23; Length 401;
Best Local Similarity 99.3%; Pred. No. 1.5e-209;
Matches 398; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MAKVLCVLYDDPVDVGVGPKTYARDLPRKIDHYPGQILPTPKAIDFTPGQLGSGVSGELG 60
Db 1 MAKVLCVLYDDPVDVGVGPKTYARDLPRKIDHYPGQILPTPKAIDFTPGQLGSGVSGELG 60
QY 61 RPYESNGHLLVYTSXDKGPDSPVERELVDADVVISQPFMPAYLTPERIAKAKMLKIALT 120
Db 61 RPYESNGHLLVYTSXDKGPDSPVERELVDADVVISQPFMPAYLTPERIAKAKMLKIALT 120
QY 121 AGTGSDDHVDOSATDRNVTVAEVTYCNSSISVAEHVMMILSLVNNYLPSEHMAKGGWNI 180
Db 121 AGTGSDDHVDOSATDRNVTVAEVTYCNSSISVAEHVMMILSLVNNYLPSEHMAKGGWNI 180
QY 181 ADCVSHAYDL EAMHVGTVAAIGLAVLRRLAPDVLHLYTDRRLPESVKEKELNLWHA 240
Db 181 ADCVSHAYDL EAMHVGTVAAIGLAVLRRLAPDVLHLYTDRRLPESVKEKELNLWHA 240
QY 241 TREDMYVCDVVTNLCP LHPETEMINDETLKLFKRGAYIVNTARGKLCORDAVARALE 300
Db 241 TREDMYVCDVVTNLCP LHPETEMINDETLKLFKRGAYIVNTARGKLCORDAVARALE 300

QY 301 GRLAGYAGDWFPQPPAPKDHPRMTPYNGMTPHISGTLTLTAQARYAAGTRILECFEGR 360
Db 301 GRLAGYAGDWFPQPPAPKDHPRMTPYNGMTPHISGTLTLTAQARYAAGTRILECFEGR 360
QY 361 PIRDEYLIYOGGALAGTGAHSYSGKNATGSGSEAAKFKKAV 401
Db 361 PIRDEYLIYOGGALAGTGAHSYSGKNATGSGSEAAKFKKAV 401
RESULT 14
AAU99107
ID AAU99107 standard; Protein: 401 AA.
XX
XX AAU99107;
XX
XX 24-SEP-2002 (first entry)
XX
XX Mycobacterium vaccae formate dehydrogenase C6S/C146S/C256S mutant.
DE
XX Formate dehydrogenase; alcohol production; ketone;
KW mutant; mutlein.
XX
XX Mycobacterium vaccae.
OS Synthetic.
OS
FH Key Location/Qualifiers
FT Misc-difference 6 /note= "Wild type Cys substituted by Ser"
FT Misc-difference 146 /note= "Wild type Cys substituted by Ser"
FT Misc-difference 256 /note= "Wild type Cys substituted by Ser"
FT
XX
XX EPI211316-A1.
XX
XX 05-JUN-2002.
XX
XX 27-NOV-2001; 2001EP-0128170.
XX
XX 29-NOV-2000; 2000JP-0363894.
XX 24-AUG-2001; 2001JP-0254631.
XX
XX (DAIIL) DAICEL CHEM IND LTD.
XX
XX Mitsuhashi K, Yamamoto H, Kimoto N;
XX
XX WPI; 2002-464925/50.
XX
XX New mutant forms of formate dehydrogenase derived from Mycobacterium
PT vaccae have enhanced activity in the presence of organic solvents and
PT are useful for producing alcohols from ketone raw materials
XX
XX
XX Claim 8; Page -: 42pp; English.
XX
XX This invention relates to a mutant form of the Mycobacterium vaccae
CC formate dehydrogenase protein which has strong formate dehydrogenase
CC activity in the presence of an organic solvent. The mutant formate
CC dehydrogenase protein of the invention may be used to produce alcohols
CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
CC proteins, the mutant polypeptides of the invention retain high enzymatic
CC activity in the presence of the alcohol product and so provides a higher
CC product yield. The present sequence represents a Mycobacterium vaccae
CC formate dehydrogenase protein mutant of the invention.
CC
XX
SQ Sequence 401 AA;
Query Match 98.6%; Score 2105; DB 23; Length 401;
Best Local Similarity 99.3%; Pred. No. 1.9e-209;
Matches 398; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MAKVLCVLYDDPVDVGVGPKTYARDLPRKIDHYPGQILPTPKAIDFTPGQLGSGVSGELG 60
Db 1 MAKVLCVLYDDPVDVGVGPKTYARDLPRKIDHYPGQILPTPKAIDFTPGQLGSGVSGELG 60

Db 1 MAKVLSVLYDDPYDGYKTYARDDLPRKIDHYPGGQILPTPKAIDFTPGQLGVSYSGLGL 60
 QY RPYLESNGHTLVVTSDDKDPDSYFEREIVADADVISOFPWPAYLTPERIKAANKLKLALT 120
 Db 61 RPYLESNGHTLVVTSDDKDPDSYFEREIVADADVISOFPWPAYLTPERIKAANKLKLALT 120
 QY 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMIISLVNRYLPSEHMARKGWT 180
 Db 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMIISLVNRYLPSEHMARKGWT 180
 QY 181 ADCVSHAYDLEAMHVGVAAGRIGLAVLRRLAPFDVLTHTDRRLPESYEKELNTWHA 240
 Db 181 ADCVSHAYDLEAMHVGVAAGRIGLAVLRRLAPFDVLTHTDRRLPESYEKELNTWHA 240
 QY 241 TREDMYPVCDDVTLNCLPLHETEMINDETLKLFKRCAYIVNTARGKLCORDAVARALE 300
 Db 241 TREDMYPVCDDVTLNCLPLHETEMINDETLKLFKRCAYIVNTARGKLCORDAVARALE 300
 QY 301 GRLAGYAGDVWFPQAPAKDHPWRTMPYNGMTPHISGTTTLAQAARYAAGTREILECFEGR 360
 Db 301 GRLAGYAGDVWFPQAPAKDHPWRTMPYNGMTPHISGTTTLAQAARYAAGTREILECFEGR 360
 QY 361 PIRDEYLIYOGGALAGTGAHSYSGNATGSEEAKEKRAV 401
 Db 361 PIRDEYLIYOGGALAGTGAHSYSGNATGSEEAKEKRAV 401

RESULT 15

AAU99120 standard; Protein; 401 AA.

24-SEP-2002 (first entry)

Mycobacterium vaccae formate dehydrogenase C65/C249S/C256V mutant.

Formate dehydrogenase; alcohol production; ketone;

mutant; mutin.

Mycobacterium vaccae.

Synthetic.

Key Location/Qualifiers

FT Misc-difference 6 /note- "Wild type Cys substituted by Ser"

FT Misc-difference 249 /note- "Wild type Cys substituted by Ser"

FT Misc-difference 256 /note- "Wild type Cys substituted by Val"

EP1211316-A1.

05-JUN-2002.

27-NOV-2001; 2001EP-0128170.

29-NOV-2000; 2000JP-0363894.

24-AUG-2001; 2001JP-0254631.

(DAIL) DAICEL CHEM IND LTD.

Mitsubishi K, Yamamoto H, Kimoto N;

WPI; 2002-464925/50.

New mutant forms of formate dehydrogenase derived from Mycobacterium

vaccae have enhanced activity in the presence of organic solvents and

are useful for producing alcohols from ketone raw materials

Example 5; Page -; 42pp; English.

This invention relates to a mutant form of the Mycobacterium vaccae

CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formated dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a Mycobacterium vaccae
 CC formate dehydrogenase protein mutant of the invention.
 XX

SQ Sequence 401 AA;

Query Match 98.6%; Score 2105; DB 23; Length 401;

Best Local Similarity .99.3%; Pred. No. 1.9e-209; Indels 0; Gaps 0;

Matches 398; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAKVLSVLYDDPYDGYKTYARDDLPRKIDHYPGGQILPTPKAIDFTPGQLGVSYSGLGL 60
 Db 1 MAKVLSVLYDDPYDGYKTYARDDLPRKIDHYPGGQILPTPKAIDFTPGQLGVSYSGLGL 60
 QY 61 RPYLESNGHTLVVTSDDKDPDSYFEREIVADADVISOFPWPAYLTPERIKAANKLKLALT 120
 Db 61 RPYLESNGHTLVVTSDDKDPDSYFEREIVADADVISOFPWPAYLTPERIKAANKLKLALT 120
 QY 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMIISLVNRYLPSEHMARKGWT 180
 Db 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMIISLVNRYLPSEHMARKGWT 180
 QY 181 ADCVSHAYDLEAMHVGVAAGRIGLAVLRRLAPFDVLTHTDRRLPESYEKELNTWHA 240
 Db 181 ADCVSHAYDLEAMHVGVAAGRIGLAVLRRLAPFDVLTHTDRRLPESYEKELNTWHA 240
 QY 241 TREDMYPVCDDVTLNCLPLHETEMINDETLKLFKRCAYIVNTARGKLCORDAVARALE 300
 Db 241 TREDMYPVCDDVTLNCLPLHETEMINDETLKLFKRCAYIVNTARGKLCORDAVARALE 300
 QY 301 GRLAGYAGDVWFPQAPAKDHPWRTMPYNGMTPHISGTTTLAQAARYAAGTREILECFEGR 360
 Db 301 GRLAGYAGDVWFPQAPAKDHPWRTMPYNGMTPHISGTTTLAQAARYAAGTREILECFEGR 360
 QY 361 PIRDEYLIYOGGALAGTGAHSYSGNATGSEEAKEKRAV 401
 Db 361 PIRDEYLIYOGGALAGTGAHSYSGNATGSEEAKEKRAV 401

Search completed: July 24, 2003, 20:46:01

Job time : 93 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:44:08 ; Search time 42 Seconds
(without alignments)
918.182 Million cell updates/sec

Title: US-09-996-008b-2

Perfect score: 2135

Sequence: 1 MAKVLCVLYDDPDVGYPKTY.....YSKGNATGSEAEAKFKKAV 401

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_76:*

1: plr1:*

2: plr2:*

3: plr3:*

4: plr4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2036	95.4	393	2 JU0334	formate dehydrogen
2	1806	84.6	400	2 UC7815	formate dehydrogen
3	1704	79.8	401	2 C95293	probable NAD-depen
4	910.5	42.6	375	2 A47117	formate dehydrogen
5	883.5	41.4	377	2 S30088	acila protein - Bme
6	872	40.8	384	2 T51423	formate dehydrogen
7	868	40.7	379	2 J02272	formate dehydrogen
8	812	38.0	364	2 UC4252	formate dehydrogen
9	803	37.6	376	2 S67300	probable membrane
10	777	36.4	374	2 E89779	NAD-dependent form
11	436	20.4	236	2 S65308	hypothetical prote
12	373	17.5	326	2 C72483	probable formate d
13	352	16.5	333	2 D75067	probable dehydroge
14	348	16.3	333	2 H69229	phosphoglycerate d
15	347	16.3	333	2 H69229	phosphoglycerate d
16	346.5	16.2	333	2 C71165	probable dehydroge
17	345.5	16.2	307	2 C71165	probable dehydroge
18	340	15.9	354	2 AB3200	phosphoglycerate d
19	338	15.8	525	1 C69705	phosphoglycerate d
20	337.5	15.8	335	2 B75057	glycerol dehydrog
21	336.5	15.8	335	2 B75057	glycerol dehydrog
22	332	15.6	323	2 E71011	probable phosphogl
23	332	15.6	328	2 H95353	2-hydroxyacid dehy
24	326	15.3	328	2 H95353	probable imported
25	323.5	15.2	524	2 A71175	phosphoglycerate d
26	319	14.9	322	2 A64437	phosphoglycerate d
27	312.5	14.6	340	2 S48189	hydroxypyruvate re
28	311.5	14.6	534	2 B83850	D-3-phosphoglycer
29	308	14.4	360	2 AB3496	gluconate 2-dehydr

30	305.5	14.3	334	2 AH2585
31	305.5	14.3	334	2 G97367
32	299	14.0	327	2 A92390
33	298.5	14.0	529	2 T35831
34	297	13.9	410	1 C64070
35	296.5	13.9	531	2 A98272
36	296.5	13.9	531	2 AF3012
37	296	13.9	318	2 G96910
38	293.5	13.7	526	2 E87647
39	292.5	13.7	533	2 A70464
40	290	13.6	527	2 E69351
41	289	13.5	334	2 G70363
42	288	13.5	318	2 AG1084
43	287	13.4	443	2 H87710
44	287	13.4	527	2 C84393
45	285.5	13.4	412	2 AD3611

ALIGNMENTS

RESULT 1

JU0334

formate dehydrogenase (EC 1.2.1.2) - Pseudomonas sp.

N:Alternate names: NAD-dependent formate dehydrogenase

C:Species: Pseudomonas sp.

C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 23-Sep-2002

C:Accession: JU0334

R:Popov, V.O.; Shumilina, I.A.; Ustinnikova, T.B.; Lamzin, V.S.; Egorov, T.A.

Bioorg. Khim. 16, 324-335, 1990

A:Title: NAD-dependent formate dehydrogenase from methylotrophic bacterium Pseudomona

A:Reference number: JU0334; MUID:90290536; PMID:2357236

A:Accession: JU0334

A:Molecule type: protein

A:Residues: 1-393 <POP>

A>Note: article in Russian with English abstract

C:Superfamily: Neurospora formate dehydrogenase

C:Keywords: homodimer; NAD; oxidoreductase

Query Match 95.4%; Score 2036; DB 2; Length 393;

Best Local Similarity 97.7%; Pred. No. 3e-161;

Matches 384; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY	2	AKVLCVLYDDPDVGYPKTYARDLPRKIDHYGGGLTPPKAIDFTPGQLGSVSGELGLR	61
DB	1	AKVLCVLYDDPDVGYPKTYARDLPRKIDHYGGGLTPPKAIDFTPGQLGSVSGELGLR	60
QY	62	PYLESNGHTLVVTSKDGSDVSFERELVDADVVISQPPMAYLTPPERIAKAKNLKALT	121
DB	61	KYLESNGHTLVVTSKDGSDVSFERELVDADVVISQPPMAYLTPPERIAKAKNLKALT	120
QY	122	GIGSDHVDLQSIDRNVTAETVYCNISVAEHVVMILSLVRNYPSEHARKGWNIA	181
DB	121	GIGSDHVDLQSIDRNVTAETVYCNISVAEHVVMILSLVRNYPSEHARKGWNIA	180
QY	182	DCVSHAYDLQSIDRNVTAETVYCNISVAEHVVMILSLVRNYPSEHARKGWNIA	241
DB	181	DCVSHAYDLQSIDRNVTAETVYCNISVAEHVVMILSLVRNYPSEHARKGWNIA	240
QY	242	REDYPCVDVTLNCPLEPTEHMINDETILFKGATVYVTAARKGLCDRAVARALESG	301
DB	241	REDYPCVDVTLNCPLEPTEHMINDETILFKGATVYVTAARKGLCDRAVARALESG	300
QY	302	RLAGYAGDVWPPQAPKDPHPRMTPYNGMTPHISGTTTLTAARVAAAGTRELCEPFG	361
DB	301	RLAGYAGDVWPPQAPKDPHPRMTPYNGMTPHISGTTTLTAARVAAAGTRELCEPFG	360
QY	362	IRDEYLLVGGALAGTGAHYSKGNATGSEAE 394	
DB	361	IRDEYLLVGGALAGTGAHYSKGNATGSEAE 393	

JC7815
 formate dehydrogenase (EC 1.2.1.2) - *Paracoccus* sp. (Strain 12-A)
 C:Species: *Paracoccus* sp.
 C:Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 03-Jun-2002
 C:Accession: JC7815
 R:Shinoda, T.; Satoh, T.; Mineki, S.; Iida, M.; Taguchi, H.
 Biosci. Biotechnol. Biochem. 66, 271-276, 2002
 A:Title: Cloning, nucleotide sequencing, and expression in *Escherichia coli* of the gene
 A:Reference number: JC7815; PMID:11999398; MUID:21994041
 A:Accession: JC7815
 A:Molecule type: DNA
 A:Residues: 1400 <SHIR>
 A:Cross-references: DDBJ:AB071373
 C:Comment: This enzyme, which catalyzes the conversion of formate to carbon dioxide with
 ti-enzyme systems, such as bioreactors involving NADH as a coenzyme, and plays a key role
 C:Genetics:
 A:Gene: fdh
 C:Keywords: oxidoreductase

Query Match 84.6%; Score 1806; DB 2; Length 400;
 Best Local Similarity 83.2%; Pred. No. 3.9e-142;
 Matches 333; Conservative 28; Mismatches 39; Indels 0; Gaps 0;

QY 1 MAKVLCVLYDDPDVGYPKTYARDDLPKIDHYPGQILPTPKAIDFTFGQLGVSSEGL 60
 D 1 MAKVAVLYDDPDVGYPTSYARDSLPIERYPDGQILPTPKAIDFVGSILGVSSEGL 60
 QY 61 RPYLESNGHTLVVTSDDKDPDSVFERELVDADVVISQFPWPAYLTPERIKAKNKLALT 120
 D 61 RNYLEAGHGLVYVTSDDKDPDSVFERELVDADVVISQFPWPAYLTPERIKAKNKLALT 120
 QY 121 AGISDHDVLDQALDRNTVTAETVYCNISVAEYVMMILSVRNYLPSHEMARKGWT 180
 D 121 AGISDHDVLDQALDRNTVTAETVYCNISVAEYVMMILSVRNYLPSHEMARKGWT 180
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRGVLAFLRPAEPDVLHTDRRLPESVEKELNTLWHA 240
 D 181 ADCVSHAYDLEAMHVGTVAAAGRGVLAFLRPAEPDVLHTDRRLPESVEKELNTLWHA 240
 QY 241 TREDMYPVCVVTINCLPHEBTEHMINDETLEKFRGAYLVNTPARKICRDVAARALE 300
 D 241 SPDMFPACDVVTINCLPHEBTEHMINDETLEKFRGAYLVNTPARKICRDVAARALE 300
 QY 301 GRLAGVAGDVPFQPAKDPHWRTPYNGMTPHISGTTLTAAQARYAAGTRELIECFEGR 360
 D 301 GOLAGVAGDVPFQPAKDPHWRTPYNGMTPHISGTTLTAAQARYAAGTRELIECFEGR 360
 QY 361 PIRDEVLYVGGALAGAGAHSGNATGSGSEAAKFKKA 400
 D 361 PIRDEVLYVGGALAGAGAHSGNATGSGSEAAKFKKA 400

RESULT 3

C95293
 probable NAD-dependent formate dehydrogenase [imported] - *Sinorhizobium meliloti* (strain
 C:Species: *Sinorhizobium meliloti*
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 23-Sep-2002
 C:Accession: C95293
 R:Barrett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
 ; Kalman, S.; Keating, D.H.; Palm, C.; Beck, M.C.; Surzycki, R.; Wells, D.H.; Yen, K.C.
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
 A:Title: Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti*
 A:Reference number: A95262; MUID:21396509; PMID:11481432
 A:Accession: C95293
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1401 <KUR>
 A:Cross-references: GB:AE006469; PIDN:AAK64909.1; PID:g14523329; GSPDB:GN00165
 A:Experimental source: strain 1021, megaplasmid pSymA
 R:Galbert, F.; Flinn, T.M.; Long, S.R.; Publer, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 peia, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
 heubert, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen,
 A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: Sma0478
 A:Genome: Plasmid
 C:Superfamily: Neurospora formate dehydrogenase

Query Match 79.8%; Score 1704; DB 2; Length 401;
 Best Local Similarity 78.4%; Pred. No. 1.2e-133;
 Matches 312; Conservative 39; Mismatches 47; Indels 0; Gaps 0;

QY 1 MAKVLCVLYDDPDVGYPKTYARDDLPKIDHYPGQILPTPKAIDFVGSILGVSSEGL 60
 D 3 MAKVAVLYDDPDVGYPTSYARDSLPIERYPDGQILPTPKAIDFVGSILGVSSEGL 62
 QY 61 RPYLESNGHTLVVTSDDKDPDSVFERELVDADVVISQFPWPAYLTPERIKAKNKLALT 120
 D 63 RKFLEGGHTLVVTSDDKDPDSVFERELVDADVVISQFPWPAYLTPERIKAKNKLALT 122
 QY 121 AGISDHDVLDQALDRNTVTAETVYCNISVAEYVMMILSVRNYLPSHEMARKGWT 180
 D 123 AGISDHDVLDQALDRNTVTAETVYCNISVAEYVMMILSVRNYLPSYQVYKGGWV 182
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRGVLAFLRPAEPDVLHTDRRLPESVEKELNTLWHA 240
 D 183 ADCVARSYDIEGHDIGVAGRGVLAFLRPAEPDVLHTDRRLPDEYAKELGYTFHQ 242
 QY 241 TREDMYPVCVVTINCLPHEBTEHMINDETLEKFRGAYLVNTPARKICRDVAARALE 300
 D 243 TAEEMVPCVVTINCLPHEBTEHMINDETLEKFRGAYLVNTPARKICRDVAARALE 302
 QY 301 GRLAGVAGDVPFQPAKDPHWRTPYNGMTPHISGTTLTAAQARYAAGTRELIECFEGR 360
 D 303 GOLAGVAGDVPFQPAKDPHWRTPYNGMTPHISGTTLTAAQARYAAGTRELIECFEGR 362
 QY 361 PIRDEVLYVGGALAGAGAHSGNATGSGSEAAKFKKA 398
 D 363 PIRDEVLYVGGALAGAGAHSGNATGSGSEAAKFKKA 400

RESULT 4

A47117
 formate dehydrogenase (EC 1.2.1.2) - *Neurospora crassa*
 C:Species: *Neurospora crassa*
 C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 23-Sep-2002
 C:Accession: A47117
 R:Chow, C.M.; RajBhandary, U.T.
 J. Bacteriol. 175, 3703-3709, 1993
 A:Title: Developmental regulation of the gene for formate dehydrogenase in *Neurospora*
 A:Reference number: A47117; MUID:93285982; PMID:8509325
 A:Accession: A47117
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-375 <CHO>
 A:Cross-references: GB:L13964; NID:g293949; PIDN:AAA99900.1; PID:g1321604
 C:Genetics:
 A:Insertions: 1/3; 17/3
 C:Superfamily: Neurospora formate dehydrogenase
 C:Keywords: oxidoreductase

Query Match 42.6%; Score 910.5; DB 2; Length 375;
 Best Local Similarity 48.4%; Pred. No. 8.0e-68;
 Matches 196; Conservative 56; Mismatches 112; Indels 41; Gaps 8;

QY 1 MAKVLCVLYDDPDVGYPKTYARDDLPKIDHYPGQILPTPKAIDFTFGQLGVSSEGL 60
 D 1 MAKVAVLYDDPDVGYPTSYARDSLPIERYPDGQILPTPKAIDFVGSILGVSSEGL 31
 QY 61 RPYLESNGHTLVVTSDDKDPDSVFERELVDADVVISQFPWPAYLTPERIKAKNKLALT 120
 D 61 RPYLESNGHTLVVTSDDKDPDSVFERELVDADVVISQFPWPAYLTPERIKAKNKLALT 120

Db 32 RKMLEDGHTLVTTCKDGENSTFDKELEDAEIIITTPHPCGYLTAERLARAKKIKLAVT 91
QY 121 AGISGDHVDLQSAIDRN--VTVAEVTCNSISVAEHVVMILSLVRNLPSPENARKGGM 178
Db 92 AGISGDHVDLNAANKNTNGITVAEYTGNSVVAEHVMTLLVLRNVPAPHEQIQEGRW 151
QY 179 NIADCVSHAYDLEAMHVTGAAGRIGLAVLRRLAPFDV-HLYTDRHRLPESVEKELNLT 237
Db 152 DVAAEAKNEFDLEGVGVGTGVRIGERVLRLKPFCKDELLYYDQGLSAKEKEICGR 211
QY 238 WHAREDMYPCVDVYVTLNCPHPEHEMINDETLKFRRGAYIVNTAGKLCDDRAVARA 297
Db 212 RVADLEEMLAQCDVYVTLNCPHPEHEMINDETLKFRRGAYIVNTAGKLCDDRAVARA 271
QY 298 LESGRLAGYAGDVWFPAPAPKDPHWR--TMPY---NGMTPHISGTTLAQARYAAGTREI 352
Db 272 LKSGHLKRGYGDVWFPAPAPKDPHWR--TMPY---NGMTPHISGTTLAQARYAAGTREI 331
QY 353 LECFEFGR-PIRDEYLIYOGGALAGTGAHYSKGNATGSEBAK 396
Db .332 IESYLGKHDYRPEDLIYGGDYA---TKSYGERERAKAAAAAAK 373

RESULT 5

S30088
acila protein - Emeritella nidulans
C:Species: Emeritella nidulans, Aspergillus nidulans
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 23-Sep-2002
C:Accession: S30088
R:Saleeba, J.A.; Cobbett, C.S.; Hynes, M.J.
Mol. Gen. Genet. 235, 349-358, 1992
A:Title: Characterization of the anda-regulated acila gene of Aspergillus nidulans.
A:Reference number: S30088; MUID:93101140; PMID:1465107
A:Accession: S30088
A:Molecule type: DNA
A:Residues: 1-377 <SAL>
A:Cross-references: EMBL:211612; NID:95544; PIDN:CAA77687.1; PID:95545
C:Genetics:
A:Gene: acila
A:Map position: 1
A:Introns: 5/3; 64/2
C:Superfamily: Neurospora formate dehydrogenase

Query Match 41.4%; Score 883.5; DB 2; Length 377;
Best Local Similarity 52.1%; Pred. No. 1.6e-65;
Matches 185; Conservative 48; Mismatches 95; Indels 27; Gaps 6;

QY 47 PGQLLSYSGELGLRPLESNGHLLVTSDDKDPDSFERELVDADVVISQFPAPYLLP 106
Db 14 PG-LTGTENELGIKRWIEQGHLLVTSDDKGENSTFDKELVDAEVIITPFGYLLA 72
QY 107 ERIKAKKLLKALTAGTSDHVDLQSAIDRN--VTVAEVTCNSISVAEHVVMILSLVR 164
Db 73 ERIKAKKLLKALTAGTSDHVDLQSAIDRN--VTVAEVTCNSISVAEHVVMILSLVR 132
QY 165 NYLPSHEWARKGNMIADCVSHAYDLEAMHVTGAAGRIGLAVLRRLAPFDV-HLYTDR 223
Db 133 NYLPSHEWARKGNMIADCVSHAYDLEAMHVTGAAGRIGLAVLRRLAPFDV-HLYTDR 192
QY 224 HRLPESVEKELNLTWHAREDMYPCVDVYVTLNCPHPEHEMINDETLKFRRGAYIV 272
Db 193 QPLRPEVEKEIGARRVDSLEEMVSCQDVYVTLNCPHPEHEMINDETLKFRRGAYIV 252
QY 273 -----LFKRGAYIVNTAGKLCDDRAVARALESRLAGYAGDVWFPAPAPKDPHWR 326
Db 253 LIIPMLAMTHKSWLVNTARALVYKEDVAEALSKSHLGGYGDVWFPAPAPKDPHWR 312
QY 327 Y-----NGMTPHISGTTLAQARYAAGTREI-PIRDEYLIYOGGALA 375
Db 313 HPWGGGNATVPHMSTSLAQIRYANGTKAILDSYFSRPFQDPQDLIVHGGDYA 367

RESULT 6

T51423
formate dehydrogenase (FDH) - Arabidopsis thaliana
N:Alternate names: protein T9L3_80
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 23-Sep-2002
C:Accession: T51423
R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.;
submitted to the Protein Sequence Database, August 2000
A:Reference number: 225394
A:Accession: T51423
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-384 <SAT>
A:Cross-references: EMBL:AL391149
A:Experimental source: cultivar Columbia; BAC clone T9L3
A:Map position: 5
A:Introns: 29/3; 88/1; 268/2; 347/3
A:Note: T9L3_80
C:Superfamily: Neurospora formate dehydrogenase

Query Match 40.8%; Score 872; DB 2; Length 384;
Best Local Similarity 52.9%; Pred. No. 1.4e-64;
Matches 172; Conservative 44; Mismatches 109; Indels 0; Gaps 0;

QY 51 LGSVGEGLRPLESNGHLLVTSDDKDPDSFERELVDADVVISQFPAPYLLP 110
Db 56 LGSVGEGLRPLESNGHLLVTSDDKDPDSFERELVDADVVISQFPAPYLLP 115
QY 111 KAKNLKLLTAGTSDHVDLQSAIDRNVTVAEYTCNSISVAEHVVMILSLVRNLPSPH 170
Db 116 KAKNLKLLTAGTSDHVDLQSAIDRNVTVAEYTCNSISVAEHVVMILSLVRNLPSPH 175
QY 171 EYARKGNMIADCVSHAYDLEAMHVTGAAGRIGLAVLRRLAPFDVHLYTDRHRLPESV 230
Db 176 NYLVGNNVAGIAYRAVDLEKGTIGYAGAGIGLTLQRLKPFCCNLLYHDLQMAEEL 235
QY 231 EKELNLTWHAREDMYPCVDVYVTLNCPHPEHEMINDETLKFRRGAYIVNTAGKLCDD 290
Db 236 EKELNLTWHAREDMYPCVDVYVTLNCPHPEHEMINDETLKFRRGAYIVNTAGKLCDD 295
QY 291 RDAVARALESRLAGYAGDVWFPAPAPKDPHWRTPYNGMTPHISGTTLAQARYAAGTR 350
Db 296 RDAVARALESRLAGYAGDVWFPAPAPKDPHWRTPYNGMTPHISGTTLAQARYAAGTR 355
QY 351 ELLECFFEGRPYRDEYLIYOGGALA 375
Db 356 DMLERYFGEDEPTEYIYKDELA 380

RESULT 7

J02272
formate dehydrogenase (EC 1.2.1.2) precursor, mitochondrial - potato
C:Species: Solanum tuberosum (potato)
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 23-Sep-2002
C:Accession: J02272; PC2189; S31325
R:Francis, Small, C.C.; Ambard-Bretteville, F.; Small, I.D.; Remy, R.
Plant Physiol. 102, 1171-1177, 1993
A:Title: Identification of a major soluble protein in mitochondria from nonphotosynth
A:Reference number: J02272; MUID:94105343; PMID:8278546
A:Accession: J02272
A:Molecule type: mRNA
A:Residues: 1-379 <DESI>
A:Cross-references: EMBL:221493; NID:q297797; PIDN:CAA79702.1; PID:q297798
A:Accession: PC2189
A:Molecule type: protein
A:Residues: 24-50 <DESI>
C:Superfamily: Neurospora formate dehydrogenase
C:Keywords: mitochondrion; oxidoreductase
F:1-23/Domain: transit peptide (mitochondrion) #status predicted <TM>
F:24-379/Product: formate dehydrogenase #status predicted <Mat>
F:199-222/Region: NAD binding #status predicted

Query Match 40.7%; Score 868; DB 2; Length 379;
Best Local Similarity 49.7%; Pred. No. 3e-64;
Matches 175; Conservative 48; Mismatches 117; Indels 12; Gaps 2;

OY 35 QILTPKAI-----DFTPGQLGSVSGELSPRYLESNGHLLVYTSDDKDPDSV 83
DB 25 QASGPKKIVGVFKANEYAEKMP-NFLGCAENALGIREWLESNGHGYIVPDKEGPCE 83
OY 84 FERELVDADVVISOPFPAVLTTPERAKAKNLKATLTAIGSDHVDLQSAIDRNVAAEV 143
DB 84 LEKIHPLHLVLTSPFPAVYTAERIKAKNLQLLTAGISGSDHDLKAAAAALITVAEV 143
OY 144 TYCSISVAEHEVVMILSLVRNYLPSHEMARKGWNIDCVSHAIVYLEAMHGVAAVARI 203
DB 144 TGSNTVSAEDELKRIILVNFLEPHGVINGEMNVAALAHRAVLEDEGKTVGVAGARI 203
OY 204 GLAVLRRLAPDPVLTDRRLPESVEKELNLVMAHTRFEDMYVVCVYVTLNCPLEHETE 263
DB 204 GRLLQRLKRPNCMLYHDLKMDSELENOIGAKFEEDLDMKSCDVIIVNTPLETKTK 263
OY 264 HMINDETLKLKRGAVYVNTARGKICDRDAVARALLESGLAGVAGDWFPOPAKDPHWR 323
DB 264 GMFKEKRIAKLKGVLLVNNARGAIMDTQAVVDACNSGHITAGISGVDWYPOPAKDPHWR 323
OY 324 TMRPNGMTPHISGTTLRQAARYAAGTRELLECFEFGRPINDEYLIYOGGALA 375
DB 324 YMPQAMTPHISGTTIDAOLRYAAGTMDRFXGDFPAENYIVXDGEIA 375

RESULT 8
JC4252
formate dehydrogenase (EC 1.2.1.2) - yeast (Candida methyllica)
C:Species: Candida methyllica
C:Date: 19-Oct-1995 #sequence_revision 08-Feb-1996 #text_change 23-Sep-2002
C:Accession: JC4252; PC4070
R:Allen, S.J.; Holbrook, J.J.
Gene 162, 99-104, 1995
A:Title: Isolation, sequence and overexpression of the gene encoding NAD-dependent formate dehydrogenase
A:Reference number: JC4252; MUID:96009885; PMID:7557425
A:Accession: JC4252
A:Molecule type: DNA
A:Residues: 1-364 <ALL>
A:References: EMBL:X81129; NID:g1181203; PIDN:CAA57036.1; PID:g1181204
A:Accession: PC4070
A:Molecule type: Protein
A:Residues: 1-30 <AL2>
A:Experimental source: ATCC 56294
C:Comment: This enzyme catalyses the final step in the dissimilatory pathway of methanol C:NAD+ to NADH.
C:Genetics:
A:Gene: fdh
C:Superfamily: Neurospora formate dehydrogenase
C:Keywords: homodimer; NAD; oxidoreductase
F:155-195/Region: beta-alpha-beta NAD nucleotide-binding fold

Query Match 38.0%; Score 812; DB 2; Length 364;
Best Local Similarity 44.6%; Pred. No. 1.3e-59;
Matches 169; Conservative 57; Mismatches 115; Indels 38; Gaps 6;

OY 3 KVLCLVLDVDDPYDGYPKTYARDLPRKIDHYEGGQLLPKKAIDTPFGQLGSVSEGLRP 62
DB 2 KIVLVLE-----AGKHADE-----EKLYGCTENKLIAN 32
OY 63 YLESNGHTLVVTSKDPDSYFERELVDADVVISOPFPAVLTTPERAKAKNLKATLTAIG 122
DB 33 WLKQOGHELTTSKDESETSELKHIRADADIIITTPHRAVITFERIDKAKNLKSVYVAG 92
OY 123 IGSDDHVDLQ--SAIDRNVAAEVTYCSISVAEHEVVMILSLVRNYLPSHEMARKGWNID 180
DB 93 VGSDDHIDLXYINQTKKISILEVTVGSNVVAEHEVVMILVAVNEFPAHEDLINHDMEV 152
OY 181 ADCVSHAYDLEAMHGVAAVGRIGLAVLRRLAPD-VHLHYTDHRLPESVEKELNLVTH 239

DB 153 AAIKADYDIEGKTIATIGAGRIGRYLERLLPENRELLYYDQALPKAEKVGARV 212
OY 240 ATREDMPVCDVYVTLNCPLEHETEMINDETLKLKRGAVYVNTARGKICDRDAVARALE 299
DB 213 ENIELVAQADIVYVNNARPLHAGTGGLINKELLSFKGAMLVNTARAKAICVAEYVAALE 272
OY 300 SGRLAGVAGDWFPOPAKDPHWRM-----PYNGMTPHISGTTLRQAARYAAGTREL 354
DB 273 SGQLRGYGDVWFPOPAKDPHWRMDMKNKYAGNAMPHYSGTTLDQTRYAEGTKNILE 332
OY 355 CFEEGR-PIREYLIYOGG 372
DB 333 SFTGKFDYRPDDITLLNG 351

RESULT 9
S67300
probable membrane protein YOR388c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein O6778
C:Species: Saccharomyces cerevisiae
C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 23-Sep-2002
C:Accession: S67300
R:Delius, H.; Hebling, U.; Hofmann, B.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67261
A:Accession: S67300
A:Molecule type: DNA
A:Residues: 1-376
A:Cross-references: EMBL:Z75296; NID:g1420834; PIDN:CAA9720.1; PID:g1420835; MIPS:YO
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:FDH1
A:Cross-references: SGD:S0005915; MIPS:YOR388C
A:Map position: 15R
C:Superfamily: Neurospora formate dehydrogenase
C:Keywords: transmembrane protein
F:121-137/Domain: transmembrane #status predicted <TM>

Query Match 37.6%; Score 803; DB 2; Length 376;
Best Local Similarity 44.7%; Pred. No. 7.5e-59;
Matches 181; Conservative 53; Mismatches 115; Indels 56; Gaps 9;

OY 3 KVLCLVLDVDDPYDGYPKTYARDLPRKIDHYEGGQLLPKKAIDTPFGQLGSVSEGLRP 62
DB 5 KVLVLVE-----GKHADEQ-----KLICIENELGIRN 35
OY 63 YLESNGHTLVVTSKDPDSYFERELVDADVVISOPFPAVLTTPERAKAKNLKATLTAIG 121
DB 36 FIEQIGELVTTIDKDEPTSYVDRELKDAEIVITTPFPAYISRRIAEAPNLKLCVTA 95
OY 122 GIGSDHVDLQSAIDRNVAAEVTYCSISVAEHEVVMILSLVRNYLPSHEMARKGWNIA 181
DB 96 GVGSDHVDLEAANRKTIVTEVGSNVVAEHEVMATILIRNYNGHQOALINGEDIA 155
OY 182 DCVSHAYDLEAMHGVAAVGRIGLAVLRRLAPD-VHLHYTDHRLPESVEKELNLVTHA 240
DB 156 GVARNEDLEDKIISTVAGRIGRYLERLVAEFPKRLIYDYOELEPAEAINRLN--EA 212
OY 241 TR-----EDMYPVCDVYVTLNCPLEHETEMINDETLKLKRGAVYVTA 285
DB 213 SKLFNGGDIYORYEKLEDMVAOSDVYVTLNCPRLKDSRGILFNKLISHMDGALVNTAR 272
OY 286 GKLCDRDAVARALLESGLAGVAGDWFPOPAKDPHWRM-----NGMTPHISGTTLR 340
DB 273 GAICVADVAEAVYSGLAGVGGVMDKQAPAKDPHWRMDMKNKYAGNAMPHYSGTTLDQTR 332
OY 341 AQARYAAGTRELLECFEGR-PIREYLIYOGGALAGTGAHSYSK 384
DB 333 AKRYAOGVNNIINSYFSKKEDYRPDDIIVQNGSYA-TRAAYGQK 376

RESULT 10
E89779


```

QY      350 RE-ILECFEGRP 361
          | | : | : |
Db      306 IENIVRVLKGKP 318

```

RESULT 13

R: anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A: Description: Pyrococcus abyssi genome sequence: Insights into archaeal chromosome structure
A: Reference number: A75001
A: Accession: D75067
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-333 <RAW>
A: Cross-references: GB:AJ248287; GB:AL096836; NID:g5458657; PIDN:CAB50433.1; PID:g545894
A: Experimental source: strain Orsay
C: Genetics:
A: Gene: PAB1008
C: Superfamily: phosphoglycerate dehydrogenase

Query Match	16.5%	Score	353;	DB 21;	Length	333;			
Best Local Similarity	32.3%	Pred. No.	1.8e-21;						
Matches	94;	Conservative	60;	Mismatches	105;	Indels	32;	Gaps	12

[illegible]

RESULT 14

phosphoglycerate dehydrogenase - Methanococcus thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 11-Jun-1999
C:Accession: H69229
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadatoro, R.; Vicalire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
K1, S.B.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
A:Reference number: A69000; MUID:8083514; PMID:9371463
A:Accession: H69229
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-525 <MTH>
A:Cross-references: GB:AE000870; GB:AE000666; NID:g2622059; PIDN:AB85466.1; PID:g262206
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH970
C:Superfamily: Bacillus phosphoglycerate dehydrogenase

Query Match 16.38; Score 348; DB 2; Length 525;

Best Local Similarity 31.7%; Pred. No. 7.6e-21;
Matches 97; Conservative 54; Mismatches 131; Indels 24; Gaps 6;

```

QY 72 WTSDDKDGPSVPERLTADAVIISQEPAPAL-----TERIRIKAK 113
   | : : : : : | : : : : : | : : : : : | : : : : : |
Db 6 VILLADSIENKGISLEEEV -AEVYVMTTTPBELLDALINDPDIAIVRSTKYTRVIEAAP 64

QY 114 NKLALTAGIGSDHVDLOSADIRNVTVAEVTYCNISVAEYVMMILSVARNYLPSEMA 173
   | : : : : : | : : : : : | : : : : : | : : : : : |
Db 65 RLKTIIRAGGVADNVVKKATDRGIMVINAPESTISVIAEHSIGMLALAKKAIARSV 124

QY 174 RKGWNIADCVSHAYDLEAMHGFTAAGRIGLAVLRRLAPDVALHYTDRRRLPESEYE 233
   | : : : : : | : : : : : | : : : : : | : : : : : |
Db 125 KEKWMEKNRFM -GIELNKGTIGIIGMERISQVYVVRKARGMDIMVYDPISEMAE -E 181

QY 234 LINTIATREDMTPVCDVYVTLNCPRLPELEHINDETLKLEKRAIYVNTARGLCDROA 293
   | : : : : : | : : : : : | : : : : : | : : : : : |
Db 182 MGVT -VTDETTLRESDIYTHVPLTPETRLHISEDERLKKDPAFYVNCARGOI IDEDA 240

QY 294 VARALESGRLAGYAGVWMPQAPARPDHWRTPYVNGMPTPISGTTLLTAQARVAAAGTREIL 353
   | : : : : : | : : : : : | : : : : : | : : : : : |
Db 241 LYRALDGEIAGALADVFEED -PGSPILLEENVLVRPHIGASTSEQRDAALIYANEI 299

QY 354 ECFEFG 359
   | : : : : : |
Db 300 KTVFQG 305

```

RESULT 15

Probable membrane protein YPL276w - Yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: hypothetical protein P0323
 C:Species: *Saccharomyces cerevisiae*
 C:Date: 10-Dec-1994 #sequence_revision 31-May-1996 #text_change 23-Sep-2002
 A:Accession: S65309; S65330
 R:Duisterhoef, A.; Floeth, M.; Filtz, M.; Hilbert, H.; Moestl, D.
 submitted to the Protein Sequence Database, May 1996
 A:Reference number: S65292
 A:Accession: S65309
 A:Molecule type: DNA
 A:Residues: 1-145 <DDE>
 A:Cross-references: EMBL:Z73633; NID:g1370566; PIDN:CAA98012.1; PID:g1370567; MIPS:YPL
 A:Experimental source: strain S288C (AB972)
 R:Delius, H.; Hebling, U
 submitted to the Protein Sequence Database, May 1996
 A:Reference number: S64967
 A:Accession: S65330
 A:Molecule type: DNA
 A:Residues: 1-145 <DED>
 A:Cross-references: EMBL:Z73633; NID:g1370566; PIDN:CAA98012.1; PID:g1370567; MIPS:YPL
 A:Experimental source: strain S288C (AB972)
 C:Genetics:
 A:Cross-references: SGD:S0006197
 A:Map position: 16L
 C:Superfamily: Neurospora formate dehydrogenase
 C:Keywords: transmembrane protein
 C:Keywords: transmembrane #status predicted <TM>
 C:121-137/Domain: transmembrane

Query Match	16.3%	Score 347;	DB 21;	Length 145;
Best Local Similarity	44.7%	Pred. No. 1.4e-21;		
Matches 76;	Conservative 27;	Mismatches 37;	Indels 30;	Gaps 3;

Qy	3 KVLCTVLYDDPVNGPYEPTAYRDLPRIDHYPGGOLIPFKALDEPFGOLISVSSEGLRP	62
Dd	5 KVLTVLYE-----GKNAHEGE-----RLTLCIENEIGIRN	35
Qy	63 YLESNHTVTYSODK-GPDVYFERELVDADAVYLSOPMPATYLPEETIAAKNKILATLP	121
	::: ::::	
Dd	36 FIEQSIGELYTTIIDKPFTSYVDNELDADELVIITPPFPYISRNNIAAPNIKLCVTA	95
Qy	122 GIGSDHVDLQSADIRNVYAETVTCNSIVAEHVMMILSLVRNYLPDSHE	171
Dd	96 GAGSDHVDLEAANERKITVTEVGTSNVVAEHMAWTLVYIRNNGQH	145

Fri Jul 25 10:38:27 2003

us-09-996-008b-2_1.rpr

Page 7

Search completed: July 24, 2003, 20:49:08
Job time : 43 secs

THIS PAGE BLANK (USPTO)

[illegible]

PT vaccae have enhanced activity in the presence of organic solvents and
 PT are useful for producing alcohols from ketone raw materials
 XX
 XX
 PS Example 1; Page 23-24; 42pp; English.

CC This invention relates to a mutant form of the *Mycobacterium vaccae*
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents the *Mycobacterium vaccae*
 CC formate dehydrogenase protein of the invention. This sequence was used
 CC to generate the mutant proteins of the invention.

XX Sequence 401 AA:

Query Match 100.0%; Score 401; DB 23; Length 401;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKVLCVLYDDPVDGPKTYARDLPRKIDHYEGGQLPPPKAIDFPGLGSSVSGEL 60
 DB 1 MAKVLCVLYDDPVDGPKTYARDLPRKIDHYEGGQLPPPKAIDFPGLGSSVSGEL 60
 QY 61 RPYLESNGHTLVVTSKDGPDSEFERELVDADVISQPPMPAYLTPPERIAKAKNKLALT 120
 DB 61 RPYLESNGHTLVVTSKDGPDSEFERELVDADVISQPPMPAYLTPPERIAKAKNKLALT 120
 QY 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVVMILSLVRNLPSEHARRKGNNI 180
 DB 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVVMILSLVRNLPSEHARRKGNNI 180
 QY 181 ADCVSHAYDLEAMHVTVAAGRIGLAVLRRLAFDYNHLYTDRHRLPESVEKELNLTWHA 240
 DB 181 ADCVSHAYDLEAMHVTVAAGRIGLAVLRRLAFDYNHLYTDRHRLPESVEKELNLTWHA 240
 QY 241 TREDMPYPCDVYTLNCPRLPETEHMINDETLLKFKGAYIVNTARSKLCDROAVARALS 300
 DB 241 TREDMPYPCDVYTLNCPRLPETEHMINDETLLKFKGAYIVNTARSKLCDROAVARALS 300
 QY 301 GRLAGYAGDVWFPQAPKDPWRTMPYNGMTPHISGTLTAAQRYAAGTRELIECFEGR 360
 DB 301 GRLAGYAGDVWFPQAPKDPWRTMPYNGMTPHISGTLTAAQRYAAGTRELIECFEGR 360
 QY 361 PIRDEYLIYOGGALAGTGAHYSKGNATGSGEAAKFKKAV 401
 DB 361 PIRDEYLIYOGGALAGTGAHYSKGNATGSGEAAKFKKAV 401

RESULT 2
 AAU99112
 ID AAU99112 standard; Protein; 401 AA.

XX AAU99112;

DT 24-SEP-2002 (first entry)

DE *Mycobacterium vaccae* formate dehydrogenase C146S mutant.

KW Formate dehydrogenase; alcohol production; ketone;
 mutant; mutein.

OS *Mycobacterium vaccae*.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 146

XX /note= "Wild type Cys substituted by Ser"

EP1211316-A1.

PD 05-JUN-2002.
 XX
 XX 27-NOV-2001; 2001EP-0128170.
 PF
 XX
 PR 29-NOV-2000; 2000JP-0363894.
 PR 24-AUG-2001; 2001JP-0254631.
 PA (DAI) DAICEL CHEM IND LTD.

PI Mitsuhashi K, Yamamoto H, Kimoto N;

DR WPI; 2002-464925/50.

PT New mutant forms of formate dehydrogenase derived from *Mycobacterium*
 PT vaccae have enhanced activity in the presence of organic solvents and
 PT are useful for producing alcohols from ketone raw materials

XX Claim 8; Page -: 42pp; English.

CC This invention relates to a mutant form of the *Mycobacterium vaccae*
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a *Mycobacterium vaccae*
 CC formate dehydrogenase protein mutant of the invention.

SQ Sequence 401 AA:

Query Match 74.8%; Score 300; DB 23; Length 401;
 Best Local Similarity 99.8%; Pred. No. 1; 1e-281;
 Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAKVLCVLYDDPVDGPKTYARDLPRKIDHYEGGQLPPPKAIDFPGLGSSVSGEL 60
 DB 1 MAKVLCVLYDDPVDGPKTYARDLPRKIDHYEGGQLPPPKAIDFPGLGSSVSGEL 60
 QY 61 RPYLESNGHTLVVTSKDGPDSEFERELVDADVISQPPMPAYLTPPERIAKAKNKLALT 120
 DB 61 RPYLESNGHTLVVTSKDGPDSEFERELVDADVISQPPMPAYLTPPERIAKAKNKLALT 120
 QY 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVVMILSLVRNLPSEHARRKGNNI 180
 DB 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVVMILSLVRNLPSEHARRKGNNI 180
 QY 181 ADCVSHAYDLEAMHVTVAAGRIGLAVLRRLAFDYNHLYTDRHRLPESVEKELNLTWHA 240
 DB 181 ADCVSHAYDLEAMHVTVAAGRIGLAVLRRLAFDYNHLYTDRHRLPESVEKELNLTWHA 240
 QY 241 TREDMPYPCDVYTLNCPRLPETEHMINDETLLKFKGAYIVNTARSKLCDROAVARALS 300
 DB 241 TREDMPYPCDVYTLNCPRLPETEHMINDETLLKFKGAYIVNTARSKLCDROAVARALS 300
 QY 301 GRLAGYAGDVWFPQAPKDPWRTMPYNGMTPHISGTLTAAQRYAAGTRELIECFEGR 360
 DB 301 GRLAGYAGDVWFPQAPKDPWRTMPYNGMTPHISGTLTAAQRYAAGTRELIECFEGR 360
 QY 361 PIRDEYLIYOGGALAGTGAHYSKGNATGSGEAAKFKKAV 401
 DB 361 PIRDEYLIYOGGALAGTGAHYSKGNATGSGEAAKFKKAV 401

RESULT 3
 AAU99113
 ID AAU99113 standard; Protein; 401 AA.

XX AAU99113;

DT 24-SEP-2002 (first entry)

DE *Mycobacterium vaccae* formate dehydrogenase C256S mutant.

XX Formate dehydrogenase; alcohol production; ketone;
 KW mutant; mutlein.
 XX
 OS Mycobacterium vaccae.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 256 /note= "Wild type Cys substituted by Ser"
 XX
 PN EPI211316-A1.
 XX
 PD 05-JUN-2002.
 XX
 PF 27-NOV-2001; 2001EP-0128170.
 XX
 XX 29-NOV-2000; 2000JP-0363894.
 PR 24-AUG-2001; 2001JP-0254631.
 XX
 PA (DAIL) DAICEL CHEM IND LTD.
 XX
 PI Mitsuhashi K, Yamamoto H, Kimoto N;
 XX
 DR WPI; 2002-464925/50.
 XX
 PT New mutant forms of formate dehydrogenase derived from Mycobacterium
 PT vaccae have enhanced activity in the presence of organic solvents and
 PT are useful for producing alcohols from ketone raw materials
 XX
 PS Claim 8; Page -: 42pp; English.
 XX
 CC This invention relates to a mutant form of the Mycobacterium vaccae
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a Mycobacterium vaccae
 CC formate dehydrogenase protein mutant of the invention.
 CC
 SQ Sequence 401 AA;

Query Match 74.8%; Score 300; DB 23; Length 401;
 Best Local Similarity 99.8%; Pred. No. 1.1e-281;
 Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAKVLCVLYDDPVGYPKTYARDLPKIDHYPGGQILPTPKAIDFTPGQLGVSSELGL 60
 DB 1 MAKVLCVLYDDPVGYPKTYARDLPKIDHYPGGQILPTPKAIDFTPGQLGVSSELGL 60
 QY 61 RPYLESNGHTLVNYSKDGPDVSFERELVDADVVISQPFMPAYLTPERIAAKNKLALT 120
 DB 61 RPYLESNGHTLVNYSKDGPDVSFERELVDADVVISQPFMPAYLTPERIAAKNKLALT 120
 QY 121 AGISGDHVDLQSAIDRNVTAEVYTCNSISVAEHVMMILSLVRNYLPSHEMARKGWN 180
 DB 121 AGISGDHVDLQSAIDRNVTAEVYTCNSISVAEHVMMILSLVRNYLPSHEMARKGWN 180
 QY 181 ADCVSHAYDLEAMHVGTVAAGRIGLAVLRRLAPFDVHLHYDRHRLPESVKEKNTLWHA 240
 DB 181 ADCVSHAYDLEAMHVGTVAAGRIGLAVLRRLAPFDVHLHYDRHRLPESVKEKNTLWHA 240
 QY 241 TREDMYPVCDDVTLNCPLEPTEHMINDETLEKFKRAYIVNTARGKICDADAARALE 300
 DB 241 TREDMYPVCDDVTLNCPLEPTEHMINDETLEKFKRAYIVNTARGKICDADAARALE 300
 QY 301 GRLAGYAGDVWFPAPAKRDMWRTMPYNGMTPHISGTTTLTAQARYAAGTRILECFEGR 360
 DB 301 GRLAGYAGDVWFPAPAKRDMWRTMPYNGMTPHISGTTTLTAQARYAAGTRILECFEGR 360
 QY 361 PIRDEYLIIVOGALAGTGAHSYSGNATGSESEAKFKKAV 401

DB 361 PIRDEYLIIVOGALAGTGAHSYSGNATGSESEAKFKKAV 401
 RESULT 4
 ID AAU99115
 ID AAU99115 standard; Protein; 401 AA.
 AC AAU99115;
 DT 24-SEP-2002 (first entry)
 DE Mycobacterium vaccae formate dehydrogenase C256v mutant.
 DE
 KW Formate dehydrogenase; alcohol production; ketone;
 KW mutant; mutlein.
 XX
 OS Mycobacterium vaccae.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 256 /note= "Wild type Cys substituted by Val"
 XX
 PN EPI211316-A1.
 XX
 PD 05-JUN-2002.
 XX
 PF 27-NOV-2001; 2001EP-0128170.
 XX
 PR 29-NOV-2000; 2000JP-0363894.
 PR 24-AUG-2001; 2001JP-0254631.
 XX
 PA (DAIL) DAICEL CHEM IND LTD.
 XX
 PI Mitsuhashi K, Yamamoto H, Kimoto N;
 XX
 DR WPI; 2002-464925/50.
 XX
 PT New mutant forms of formate dehydrogenase derived from Mycobacterium
 PT vaccae have enhanced activity in the presence of organic solvents and
 PT are useful for producing alcohols from ketone raw materials
 XX
 PS Claim 8; Page -: 42pp; English.
 XX
 CC This invention relates to a mutant form of the Mycobacterium vaccae
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a Mycobacterium vaccae
 CC formate dehydrogenase protein mutant of the invention.
 CC
 SQ Sequence 401 AA;

Query Match 74.8%; Score 300; DB 23; Length 401;
 Best Local Similarity 99.8%; Pred. No. 1.1e-281;
 Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAKVLCVLYDDPVGYPKTYARDLPKIDHYPGGQILPTPKAIDFTPGQLGVSSELGL 60
 DB 1 MAKVLCVLYDDPVGYPKTYARDLPKIDHYPGGQILPTPKAIDFTPGQLGVSSELGL 60
 QY 61 RPYLESNGHTLVNYSKDGPDVSFERELVDADVVISQPFMPAYLTPERIAAKNKLALT 120
 DB 61 RPYLESNGHTLVNYSKDGPDVSFERELVDADVVISQPFMPAYLTPERIAAKNKLALT 120
 QY 121 AGISGDHVDLQSAIDRNVTAEVYTCNSISVAEHVMMILSLVRNYLPSHEMARKGWN 180
 DB 121 AGISGDHVDLQSAIDRNVTAEVYTCNSISVAEHVMMILSLVRNYLPSHEMARKGWN 180

PS Claim 8; Page -: 42pp; English.

CC This invention relates to a mutant form of the *Mycobacterium vaccae*
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a *Mycobacterium vaccae*
 CC formate dehydrogenase protein mutant of the invention.

XX Sequence 401 AA;

Query Match 73.3%; Score 294; DB 23; Length 401;

Best Local Similarity 99.7%; Pred. No. 6.8e-276;

Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 VLVDPPVGYPKTYARDDLPKIDHYPGQILPTPKAIDFTPGQLGSVSGELGRPYLES 66
 Db 7 VLDDPPVGYPKTYARDDLPKIDHYPGQILPTPKAIDFTPGQLGSVSGELGRPYLES 66
 QY 67 NGHTLVVTSKDGDPSVFERELVDADYVISQPFMPAYLTPERIAKANLKLALTAIGISD 126
 Db 67 NGHTLVVTSKDGDPSVFERELVDADYVISQPFMPAYLTPERIAKANLKLALTAIGISD 126
 QY 127 HVDLQSAIDRNVTAAEVTYCNISVAEHVMMILSLVKNYLPSEHMARKGGMNTADCVSH 186
 Db 127 HVDLQSAIDRNVTAAEVTYCNISVAEHVMMILSLVKNYLPSEHMARKGGMNTADCVSH 186
 QY 187 AYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDRHRLPESVEKELNTWHAATREDMY 246
 Db 187 AYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDRHRLPESVEKELNTWHAATREDMY 246
 QY 247 PVCDDVTLNCPRLHETEHMINDETLKLFRKRAYIVNTARGKLCDDAARALLESGRLAGY 306
 Db 247 PVCDDVTLNCPRLHETEHMINDETLKLFRKRAYIVNTARGKLCDDAARALLESGRLAGY 306
 QY 307 AGDVWFPOPAPKDPHWRMPYNGMTPHISGTTLTAAQRYAAGTRELIECFEGRPIRDY 366
 Db 307 AGDVWFPOPAPKDPHWRMPYNGMTPHISGTTLTAAQRYAAGTRELIECFEGRPIRDY 366
 QY 367 LIVOGALAGTGAHSYSKGNATGSEEAKEFKKAY 401
 Db 367 LIVOGALAGTGAHSYSKGNATGSEEAKEFKKAY 401

RESULT 7
 AAU99110 standard; Protein: 401 AA.

AC AAU99110;

DT 24-SEP-2002 (first entry)

DE *Mycobacterium vaccae* formate dehydrogenase C65/C256a mutant.

KW Formate dehydrogenase; alcohol production; ketone;

KM mutant; mutlein.

OS *Mycobacterium vaccae*.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 6 /note= "Wild type Cys substituted by Ser"

FT Misc-difference 256 /note= "Wild type Cys substituted by Ser"

FT /note= "Wild type Cys substituted by Ala"

PN EPI211316-A1.

PD 05-JUN-2002.

XX

PF 27-NOV-2001; 2001EP-0128170.

XX 29-NOV-2000; 2000JP-0363894.

PR 24-AUG-2001; 2001JP-0254631.

XX (DAI) DAICEL CHEM IND LTD.

PA Mitsunashi K, Yamamoto H, Kimoto N;

PI WPI; 2002-464925/50.

XX New mutant forms of formate dehydrogenase derived from *Mycobacterium*

PT *vaccae* have enhanced activity in the presence of organic solvents and

PT are useful for producing alcohols from ketone raw materials

XX Claim 8; Page -: 42pp; English.

CC This invention relates to a mutant form of the *Mycobacterium vaccae*
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a *Mycobacterium vaccae*
 CC formate dehydrogenase protein mutant of the invention.

XX Sequence 401 AA;

Query Match 73.3%; Score 294; DB 23; Length 401;

Best Local Similarity 99.7%; Pred. No. 6.8e-276;

Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 VLVDPPVGYPKTYARDDLPKIDHYPGQILPTPKAIDFTPGQLGSVSGELGRPYLES 66
 Db 7 VLDDPPVGYPKTYARDDLPKIDHYPGQILPTPKAIDFTPGQLGSVSGELGRPYLES 66
 QY 67 NGHTLVVTSKDGDPSVFERELVDADYVISQPFMPAYLTPERIAKANLKLALTAIGISD 126
 Db 67 NGHTLVVTSKDGDPSVFERELVDADYVISQPFMPAYLTPERIAKANLKLALTAIGISD 126
 QY 127 HVDLQSAIDRNVTAAEVTYCNISVAEHVMMILSLVKNYLPSEHMARKGGMNTADCVSH 186
 Db 127 HVDLQSAIDRNVTAAEVTYCNISVAEHVMMILSLVKNYLPSEHMARKGGMNTADCVSH 186
 QY 187 AYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDRHRLPESVEKELNTWHAATREDMY 246
 Db 187 AYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDRHRLPESVEKELNTWHAATREDMY 246
 QY 247 PVCDDVTLNCPRLHETEHMINDETLKLFRKRAYIVNTARGKLCDDAARALLESGRLAGY 306
 Db 247 PVCDDVTLNCPRLHETEHMINDETLKLFRKRAYIVNTARGKLCDDAARALLESGRLAGY 306
 QY 307 AGDVWFPOPAPKDPHWRMPYNGMTPHISGTTLTAAQRYAAGTRELIECFEGRPIRDY 366
 Db 307 AGDVWFPOPAPKDPHWRMPYNGMTPHISGTTLTAAQRYAAGTRELIECFEGRPIRDY 366
 QY 367 LIVOGALAGTGAHSYSKGNATGSEEAKEFKKAY 401
 Db 367 LIVOGALAGTGAHSYSKGNATGSEEAKEFKKAY 401

RESULT 8
 AAU99111 standard; Protein: 401 AA.

AC AAU99111;

DT 24-SEP-2002 (first entry)

DE *Mycobacterium vaccae* formate dehydrogenase C65/C256v mutant.

KW Formate dehydrogenase; alcohol production; ketone;

XX

Key	Location/Qualifiers
Key	Location/Qualifiers
Misc-difference 6	/note= "Wild type Cys substituted by Ser"
Misc-difference 256	/note= "Wild type Cys substituted by Val"
Ep1211316-A1.	
05-JUN-2002.	
27-NOV-2001; 2001EP-0128170.	
29-NOV-2000; 2000JP-0363894.	
24-AUG-2001; 2001JP-0254631.	
(DAICEL CHEM IND LTD.	
Mitsubashi K, Yamamoto H, Kimoto N;	
WPI; 2002-464925/50.	
New mutant forms of formate dehydrogenase derived from Mycobacterium vaccae have enhanced activity in the presence of organic solvents and are useful for producing alcohols from ketone raw materials	
Claim 8; Page -: 42pp; English.	
This invention relates to a mutant form of the Mycobacterium vaccae formate dehydrogenase protein which has strong formate dehydrogenase activity in the presence of an organic solvent. The mutant formate dehydrogenase protein of the invention may be used to produce alcohols using ketones as raw materials. Unlike prior formatted dehydrogenase, proteins, the mutant polypeptides of the invention retain high enzymatic activity in the presence of the alcohol product and so provides a higher product yield. The present sequence represents a Mycobacterium vaccae formate dehydrogenase protein mutant of the invention.	
Sequence 401 AA;	
Query Match 73.3%; Score 294; DB 23; Length 401;	
Best Local Similarity 99.7%; Pred. No. 6.8e-276;	
Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
7 VLYDDPVGVGYRKYRADDLPKIDHYPGGQILLPTPKAIDFTPGQLGSGVSGELGIRPYLES 66	
7 VLYDDPVGVGYRKYRADDLPKIDHYPGGQILLPTPKAIDFTPGQLGSGVSGELGIRPYLES 66	
67 NGHTLVATSDKDGDPDSVERELVDADVVISQPFWPAYLTPERIAKAKMILKIALTAGIGSD 126	
67 NGHTLVATSDKDGDPDSVERELVDADVVISQPFWPAYLTPERIAKAKMILKIALTAGIGSD 126	
127 HVDLOSADRNVATAEVTYCNSSISVAEHVMMILSVANNYLPESHMARKGWNITADCVSH 186	
127 HVDLOSADRNVATAEVTYCNSSISVAEHVMMILSVANNYLPESHMARKGWNITADCVSH 186	
187 AYDLEAMHVGIVAAGRIGLAVLRLAPDPVHLHTYDRRLRLESPYKEKINTLWHTATREDMY 246	
187 AYDLEAMHVGIVAAGRIGLAVLRLAPDPVHLHTYDRRLRLESPYKEKINTLWHTATREDMY 246	
247 PVCADVTLTNCPLHPEHINDETLLKLEKRGAYIYNTARGKLCDRDAVARALESGRLAGY 306	
247 PVCADVTLTNCPLHPEHINDETLLKLEKRGAYIYNTARGKLCDRDAVARALESGRLAGY 306	
307 AGDWVFPQPAKRDHWRTPYNGMTPHISGTTLTAAQARYAAGTRILECFEFGRIIDEX 366	
307 AGDWVFPQPAKRDHWRTPYNGMTPHISGTTLTAAQARYAAGTRILECFEFGRIIDEX 366	
367 LIYOGALAGTGAHSYSGKNTGGSEEAARFKKAV 401	

```

DB      367  LIYOGSALAGTGAHSYSGKNTGSGSEAAKTKKAV 401
|||||
RESULT 9
AAU99117
ID      AAU99117 standard; Protein; 401 AA.
XX
XX      AAU99117;
XX
XX      24-SEP-2002 (first entry)
XX
XX      Mycobacterium vaccae formate dehydrogenase C6A/C256V mutant.
DE
XX      Formate dehydrogenase; alcohol production; ketone;
XX      mutant; muteln.
XX
XX      Mycobacterium vaccae.
OS      Synthetic.
XX
XX      Key      Location/Qualifiers
FH      Misc-difference 6 /note= "Wild type Cys substituted by Ala"
FT      Misc-difference 256 /note= "Wild type Cys substituted by Val"
FT      Misc-difference 256 /note= "Wild type Cys substituted by Val"
XX
XX      EPI211316-AL.
XX
XX      05-JUN-2002.
XX
XX      27-NOV-2001; 2001EP-0128170.
XX
XX      29-NOV-2000; 2000JP-0363894.
XX      PR      24-AUG-2001; 2001JP-0254631.
XX
XX      (DAIL ) DAICEL CHEM IND LTD.
XX
XX      Mitsuhashi K, Yamamoto H, Kimoto N;
XX
XX      WPI; 2002-464925/50.
XX
XX      New mutant forms of formate dehydrogenase derived from Mycobacterium
XX      vaccae have enhanced activity in the presence of organic solvents and
XX      are useful for producing alcohols from ketone raw materials
XX
XX      Claim 8; Page -: 42pp; English.
XX
XX      This invention relates to a mutant form of the Mycobacterium vaccae
XX      formate dehydrogenase protein which has strong formate dehydrogenase
XX      activity in the presence of an organic solvent. The mutant formate
XX      dehydrogenase protein of the invention may be used to produce alcohols
XX      using ketones as raw materials. Unlike prior formatted dehydrogenase,
XX      proteins, the mutant polypeptides of the invention retain high enzymatic
XX      activity in the presence of the alcohol product and so provides a higher
XX      product yield. The present sequence represents a Mycobacterium vaccae
XX      formate dehydrogenase protein mutant of the invention.
XX
XX      Sequence      401 AA:
XX
XX      Query Match      73.3%; Score 294; DB 23; Length 401;
XX      Best Local Similarity 99.7%; Pred. No. 6.8e-276;
XX      Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY      7 VLYDDPVNDGVPKRYARDDLPKIDHYPGQILPTPKAIDFTPGQLGVSYGELGRPYLES 66
OY      |||||||
OY      7 VLYDDPVNDGVPKRYARDDLPKIDHYPGQILPTPKAIDFTPGQLGVSYGELGRPYLES 66
OY      |||||||
OY      67 NGHLTVYTSKDDGSDPSVFERELVADAVYISQPFMPAYLTTPRIRAKKALKALTAAGTSD 126
OY      |||||||
OY      67 NGHLTVYTSKDDGSDPSVFERELVADAVYISQPFMPAYLTTPRIRAKKALKALTAAGTSD 126
OY      |||||||
OY      127 HVLDLSADRVTVAEVTYCNISIVAEHVMMILVNTYLPSEHMAKKGMMTADCVSH 186
OY      |||||||

```

```
Db 127 HVDLSAIDRNVTAEVITYCNSISVAEHVMMILSLVNYLPSEHMARKGMNIADCVSH 186
OY 187 AYDLEAMHVGTVAAGRIGLAVLRRLAPFDVHLHYTDHRHRLPESYEKELNTWHAATREDMY 246
Db 187 AYDLEAMHVGTVAAGRIGLAVLRRLAPFDVHLHYTDHRHRLPESYEKELNTWHAATREDMY 246
OY 247 PVCVVTLNCLHRETEHIMIDETLKLFRKCAIYVYNARGLCDRDAAVAALESGLAGY 306
Db 247 PVCVVTLNCLHRETEHIMIDETLKLFRKCAIYVYNARGLCDRDAAVAALESGLAGY 306
OY 307 AGDVWFPQAPKDHPRMTMPNGMTPHISGTTTLTAQARYAAGTREILCEFEGRPIRDEY 366
Db 307 AGDVWFPQAPKDHPRMTMPNGMTPHISGTTTLTAQARYAAGTREILCEFEGRPIRDEY 366
OY 367 LIVOGGALAGTGAHSHYSKGNATGSGSEAKFKKAV 401
Db 367 LIVOGGALAGTGAHSHYSKGNATGSGSEAKFKKAV 401

RESULT 10
AU09121
ID AU09121 standard; Protein: 401 AA.
AC AU09121;
XX
XX 24-SEP-2002 (first entry)
DE Mycobacterium vaccae formate dehydrogenase C65/C256V/C355S mutant.
KW Formate dehydrogenase; alcohol production; ketone;
KW mutant; muteln.
XX
XX Mycobacterium vaccae.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 6 /note= "Wild type Cys substituted by Ser"
FT Misc-difference 256 /note= "Wild type Cys substituted by Val"
FT Misc-difference 256 /note= "Wild type Cys substituted by Val"
FT Misc-difference 256 /note= "Wild type Cys substituted by Ser"
XX
XX EPI211316-A1.
XX
XX 05-JUN-2002.
XX
XX 27-NOV-2001; 2001EP-0128170.
XX
XX 29-NOV-2000; 2000JP-0363894.
XX
XX 24-AUG-2001; 2001JP-0254631.
XX
XX (DAIL ) DAICEL CHEM IND LTD.
XX
XX Mitsunashi K, Yamamoto H, Kimoto N;
XX
XX WPI: 2002-464925/50.
XX
XX New mutant forms of formate dehydrogenase derived from Mycobacterium
XX vaccae have enhanced activity in the presence of organic solvents and
XX are useful for producing alcohols from ketone raw materials
XX
XX Example 6; Page -: 42pp; English.
XX
XX This invention relates to a mutant form of the Mycobacterium vaccae
XX formate dehydrogenase protein which has strong formate dehydrogenase
XX activity in the presence of an organic solvent. The mutant formate
XX dehydrogenase protein of the invention may be used to produce alcohols
XX using ketones as raw materials. Unlike prior formatted dehydrogenase,
XX proteins, the mutant polypeptides of the invention retain high enzymatic
XX activity in the presence of the alcohol product and so provides a higher
XX product yield. The present sequence represents a Mycobacterium vaccae
XX formate dehydrogenase protein mutant of the invention.
```

```
XX SQ Sequence 401 AA;
XX Query Match 62.1%; Score 249; DB 23; Length 401;
XX Best Local Similarity 100.0%; Pred. No. 2,Se-232;
XX Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 VLYDDPYDGYKRYTARDDLPKIDHYPGQILPTPKAIDFTPGQLGSVSGELGRPYLES 66
Db 7 VLYDDPYDGYKRYTARDDLPKIDHYPGQILPTPKAIDFTPGQLGSVSGELGRPYLES 66
OY 67 NGHTLVYSDSDGDSVFERELVDADVVISQPFMPAYLTPEIRIAKAKLKALTAAGIGSD 126
Db 67 NGHTLVYSDSDGDSVFERELVDADVVISQPFMPAYLTPEIRIAKAKLKALTAAGIGSD 126
OY 127 HVDLSAIDRNVTAEVITYCNSISVAEHVMMILSLVNYLPSEHMARKGMNIADCVSH 186
Db 127 HVDLSAIDRNVTAEVITYCNSISVAEHVMMILSLVNYLPSEHMARKGMNIADCVSH 186
OY 187 AYDLEAMHVGTVAAGRIGLAVLRRLAPFDVHLHYTDHRHRLPESYEKELNTWHAATREDMY 246
Db 187 AYDLEAMHVGTVAAGRIGLAVLRRLAPFDVHLHYTDHRHRLPESYEKELNTWHAATREDMY 246
OY 247 PVCVVTLN 255
Db 247 PVCVVTLN 255

RESULT 11
AU09120
ID AU09120 standard; Protein: 401 AA.
AC AU09120;
XX
XX 24-SEP-2002 (first entry)
DE Mycobacterium vaccae formate dehydrogenase C65/C249S/C256V mutant.
KW Formate dehydrogenase; alcohol production; ketone;
KW mutant; muteln.
XX
XX Mycobacterium vaccae.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 6 /note= "Wild type Cys substituted by Ser"
FT Misc-difference 249 /note= "Wild type Cys substituted by Ser"
FT Misc-difference 256 /note= "Wild type Cys substituted by Val"
FT Misc-difference 256 /note= "Wild type Cys substituted by Val"
XX
XX EPI211316-A1.
XX
XX 05-JUN-2002.
XX
XX 27-NOV-2001; 2001EP-0128170.
XX
XX 29-NOV-2000; 2000JP-0363894.
XX
XX 24-AUG-2001; 2001JP-0254631.
XX
XX (DAIL ) DAICEL CHEM IND LTD.
XX
XX Mitsunashi K, Yamamoto H, Kimoto N;
XX
XX WPI: 2002-464925/50.
XX
XX New mutant forms of formate dehydrogenase derived from Mycobacterium
XX vaccae have enhanced activity in the presence of organic solvents and
XX are useful for producing alcohols from ketone raw materials
XX
XX Example 5; Page -: 42pp; English.
XX
```

CC This invention relates to a mutant form of the Mycobacterium vaccae
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a Mycobacterium vaccae
 CC formate dehydrogenase protein mutant of the invention.

SO Sequence 401 AA;

Query Match

Best Local Similarity 100.0%; Score 242; DB 23; Length 401;
 Pred. No. 1.5e-225;

Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 VLYDDVDGPKRYADDDLPKIDHYGQTLPPPKAIDFPGLLSVSGELRLPYLES 66
 Db 7 VLYDDVDGPKRYADDDLPKIDHYGQTLPPPKAIDFPGLLSVSGELRLPYLES 66

QY 67 NGHTLVTSDDKSPDSVFERELVDADVISQPFMPAYLPPERTAKAKNLKALTAGIGSD 126
 Db 67 NGHTLVTSDDKSPDSVFERELVDADVISQPFMPAYLPPERTAKAKNLKALTAGIGSD 126

QY 127 HVDLQSAIDRNVTVAEYTCNSISVAEHVVMILSLVRNYLPSEHARKGWNIIADCVSH 186
 Db 127 HVDLQSAIDRNVTVAEYTCNSISVAEHVVMILSLVRNYLPSEHARKGWNIIADCVSH 186

QY 187 AYLLEMHVGTVAAGRIGLAVLRRLAPFDVHLATYTRHRRLPEVEKEELNTWTATREDMY 246
 Db 187 AYLLEMHVGTVAAGRIGLAVLRRLAPFDVHLATYTRHRRLPEVEKEELNTWTATREDMY 246

QY 247 PV 248
 Db 247 PV 248

RESULT 12
 AAM37857
 ID AAM37857 standard; Protein; 401 AA.

AC AAM37857;

DT 10-AUG-1998 (first entry)

DE Formate dehydrogenase.

KM Formate dehydrogenase; recombinant plasmid; (NAD⁺)-dependent;

OS NADH-dependent.

OS Pseudomonas sp.

OS Pseudomonas sp.

FT Key Location/Qualifiers
 FT Misc-difference 31..58 "in this amino acid sequence, residues 49-58
 FT /note- appear to have been mis-inserted, and should
 FT be in position 31"

FT Misc-difference 151..178 "in this amino acid sequence, residues 169-178
 FT /note- appear to have been mis-inserted, and should
 FT be in position 151"

FT JP10023896-A.

PN 27-JAN-1998.

PD 19-AUG-1996; 96JP-0217060.

PR 07-MAY-1996; 96JP-0112303.

XX (NIRA) UNITIKA LTD.

XX WPI; 1998-152798/14.

DR N-PSDB; AAV29010.

XX New recombinant plasmid - comprises genes encoding NAD⁺ dependent
 PT formate (sic) dehydrogenase and NAD dependent amino acid
 PT dehydrogenase

PS Example 1; Pages 9-10; 13pp; Japanese.

CC This is the amino acid sequence of formate dehydrogenase. It is
 CC used in the method of the invention to create new recombinant
 CC plasmid comprising a gene encoding a NAD⁺ dependent formate
 CC dehydrogenase, and a gene encoding, an NADH-dependent amino acid
 CC dehydrogenase.

SO Sequence 401 AA;

Query Match

Best Local Similarity 55.6%; Score 223; DB 19; Length 401;
 Pred. No. 3.7e-207;

Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 NINDCVSHAYDLEAMHVGTAAAGRIGLAVLRRLAPFDVHLATYTRHRRLPEVEKEELNTWT 238
 Db 179 NINDCVSHAYDLEAMHVGTAAAGRIGLAVLRRLAPFDVHLATYTRHRRLPEVEKEELNTWT 238

QY 239 HATREDMPYCDVVTNLCPLEPETEMINDETLEKRGAYIVNTARGKICDRDAVARAL 298
 Db 239 HATREDMPYCDVVTNLCPLEPETEMINDETLEKRGAYIVNTARGKICDRDAVARAL 298

QY 299 ESGRLAGYAGDWFPPAPARDHWRTPNGMTPHISGTTLTAQARYAAGTRILECFEE 358
 Db 299 ESGRLAGYAGDWFPPAPARDHWRTPNGMTPHISGTTLTAQARYAAGTRILECFEE 358

QY 359 GRPIRDEYLIYOGGALAGTAGHSHSKGNATGSGSEAAKFKKAY 401
 Db 359 GRPIRDEYLIYOGGALAGTAGHSHSKGNATGSGSEAAKFKKAY 401

RESULT 13
 AAU99114
 ID AAU99114 standard; Protein; 401 AA.

AC AAU99114;

DT 24-SEP-2002 (first entry)

DE Mycobacterium vaccae formate dehydrogenase C146S/C256S mutant.

KM Formate dehydrogenase; alcohol production; ketone;
 mutant; mutain.

OS Mycobacterium vaccae.

OS Synthetic.

FT Key Location/Qualifiers
 FT Misc-difference 146 "wild type Cys substituted by Ser"
 FT /note- "wild type Cys substituted by Ser"
 FT Misc-difference 256 "wild type Cys substituted by Ser"

FT EP1211316-A1.

PN 05-JUN-2002.

PD 27-NOV-2001; 2001EP-0128170.

PF 29-NOV-2000; 2000JP-0363894.

PR 24-AUG-2001; 2001JP-0254631.

XX (DAIL) DAICEL CHEM IND LTD.

XX Mitsubishi K, Yamamoto H, Kimoto N;

XX WPI; 2002-464925/50.

XX New mutant forms of formate dehydrogenase derived from *Mycobacterium*
 PT vaccae have enhanced activity in the presence of organic solvents and
 PT are useful for producing alcohols from ketone raw materials -
 XX
 XX Claim 8; Page -: 42pp; English.
 CC This invention relates to a mutant form of the *Mycobacterium vaccae*
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a *Mycobacterium vaccae*
 CC formate dehydrogenase protein mutant of the invention.
 CC
 SQ Sequence 401 AA;
 XX
 Query Match 49.6%; Score 199; DB 23; Length 401;
 Best Local Similarity 99.5%; Pred. No. 6.3e-184;
 Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MAKVLCVLYDDPVDGPKRYARDDLPKIDHYPGQIIPTRKAIDFTPGQLGSVSGELGL 60
 DB 1 MAKVLCVLYDDPVDGPKRYARDDLPKIDHYPGQIIPTRKAIDFTPGQLGSVSGELGL 60
 QY 61 RPYLESNGHTLVVTSDDKGPDSVFERELVDADVVISQPFMPAYLTPERIAKAKNLKALT 120
 DB 61 RPYLESNGHTLVVTSDDKGPDSVFERELVDADVVISQPFMPAYLTPERIAKAKNLKALT 120
 QY 121 AGISGDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMILSLVNNYLPSEHMARKGWINI 180
 DB 121 AGISGDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMILSLVNNYLPSEHMARKGWINI 180
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPPDVHLHTDRRLPESVEKELNLTWHA 240
 DB 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPPDVHLHTDRRLPESVEKELNLTWHA 240
 QY 241 TREDMYPVCQVVTNCLPHPETEHMINDETLTKRKRAYIVNTARGKLCDDAVARALES 300
 DB 241 TREDMYPVCQVVTNCLPHPETEHMINDETLTKRKRAYIVNTARGKLCDDAVARALES 300
 QY 301 GRLAGYAGDWFPPAPKRDHPWRTMPYNGMTPHISGTTLTQAQARYAAGTREILECFEGR 360
 DB 301 GRLAGYAGDWFPPAPKRDHPWRTMPYNGMTPHISGTTLTQAQARYAAGTREILECFEGR 360
 QY 361 PIRDEYLIYOGALAGTAGHASYSKGNATGSGESEAARFKKAV 401
 DB 361 PIRDEYLIYOGALAGTAGHASYSKGNATGSGESEAARFKKAV 401
 RESULT 14
 AAU09116 standard; Protein: 401 AA.
 XX
 AC AAU09116;
 XX
 XX 24-SEP-2002 (first entry)
 XX
 DE *Mycobacterium vaccae* formate dehydrogenase C146S/C256V mutant.
 XX
 KW Formate dehydrogenase; alcohol production; ketone;
 KM mutant; muten.
 XX
 OS *Mycobacterium vaccae*.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Msc-difference 146 /note- "Wild type Cys substituted by Ser"
 FT Msc-difference 256 /note- "Wild type Cys substituted by Val"
 FT

XX
 PN EPI211316-A1.
 XX
 PD 05-JUN-2002.
 XX
 XX 27-NOV-2001; 2001EP-0128170.
 PE
 XX 29-NOV-2000; 2000JP-0363894.
 PR 24-AUG-2001; 2001JP-0254631.
 PA (DAIL) DAICEL CHEM IND LTD.
 PI Mitsunashi K, Yamamoto H, Kimoto N;
 XX WPI: 2002-464925/50.
 DR
 XX
 PT New mutant forms of formate dehydrogenase derived from *Mycobacterium*
 PT vaccae have enhanced activity in the presence of organic solvents and
 PT are useful for producing alcohols from ketone raw materials -
 XX
 XX Claim 8; Page -: 42pp; English.
 CC This invention relates to a mutant form of the *Mycobacterium vaccae*
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a *Mycobacterium vaccae*
 CC formate dehydrogenase protein mutant of the invention.
 CC
 SQ Sequence 401 AA;
 XX
 Query Match 49.6%; Score 199; DB 23; Length 401;
 Best Local Similarity 99.5%; Pred. No. 6.3e-184;
 Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MAKVLCVLYDDPVDGPKRYARDDLPKIDHYPGQIIPTRKAIDFTPGQLGSVSGELGL 60
 DB 1 MAKVLCVLYDDPVDGPKRYARDDLPKIDHYPGQIIPTRKAIDFTPGQLGSVSGELGL 60
 QY 61 RPYLESNGHTLVVTSDDKGPDSVFERELVDADVVISQPFMPAYLTPERIAKAKNLKALT 120
 DB 61 RPYLESNGHTLVVTSDDKGPDSVFERELVDADVVISQPFMPAYLTPERIAKAKNLKALT 120
 QY 121 AGISGDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMILSLVNNYLPSEHMARKGWINI 180
 DB 121 AGISGDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMILSLVNNYLPSEHMARKGWINI 180
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPPDVHLHTDRRLPESVEKELNLTWHA 240
 DB 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPPDVHLHTDRRLPESVEKELNLTWHA 240
 QY 241 TREDMYPVCQVVTNCLPHPETEHMINDETLTKRKRAYIVNTARGKLCDDAVARALES 300
 DB 241 TREDMYPVCQVVTNCLPHPETEHMINDETLTKRKRAYIVNTARGKLCDDAVARALES 300
 QY 301 GRLAGYAGDWFPPAPKRDHPWRTMPYNGMTPHISGTTLTQAQARYAAGTREILECFEGR 360
 DB 301 GRLAGYAGDWFPPAPKRDHPWRTMPYNGMTPHISGTTLTQAQARYAAGTREILECFEGR 360
 QY 361 PIRDEYLIYOGALAGTAGHASYSKGNATGSGESEAARFKKAV 401
 DB 361 PIRDEYLIYOGALAGTAGHASYSKGNATGSGESEAARFKKAV 401
 RESULT 15
 AAU09107 standard; Protein: 401 AA.
 XX
 AC AAU09107;
 XX

DT 24-SEP-2002 (first entry)
XX Mycobacterium vaccae formate dehydrogenase C6S/C146S/C256S mutant.
DE Mycobacterium vaccae formate dehydrogenase; alcohol production; ketone;
XX Formate dehydrogenase; alcohol production; ketone;
KW mutant; muteln.
XX Mycobacterium vaccae.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT MISC-difference 6 /note= "Wild type Cys substituted by Ser"
FT MISC-difference 146 /note= "Wild type Cys substituted by Ser"
FT MISC-difference 256 /note= "Wild type Cys substituted by Ser"
FT MISC-difference 256 /note= "Wild type Cys substituted by Ser"
XX
PN EP1211316-A1.
XX
XX 05-JUN-2002.
PD
XX 27-NOV-2001; 2001EP-0128170.
PF
XX 29-NOV-2000; 2000JP-0363894.
PR 24-AUG-2001; 2001JP-0254631.
XX
XX (DAIL) DAICEL CHEM IND LTD.
PA
XX Mitsubashi K, Yamamoto H, Kimoto N;
PI WPI; 2002-464925/50.
DR
XX
XX New mutant forms of formate dehydrogenase derived from Mycobacterium
PT vaccae have enhanced activity in the presence of organic solvents and
PT are useful for producing alcohols from ketone raw materials
XX
PS Claim 8; Page -: 42pp; English.
XX
CC This invention relates to a mutant form of the Mycobacterium vaccae
CC formate dehydrogenase protein which has strong formate dehydrogenase
CC activity in the presence of an organic solvent. The mutant formate
CC dehydrogenase protein of the invention may be used to produce alcohols
CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
CC proteins, the mutant polypeptides of the invention retain high enzymatic
CC activity in the presence of the alcohol product and so provides a higher
CC product yield. The present sequence represents a Mycobacterium vaccae
CC formate dehydrogenase protein mutant of the invention.
XX
SQ Sequence 401 AA:

Query Match 48.1%; Score 193; DB 23; Length 401;
Best Local Similarity 99.5%; Pred. No. 4.1e-178;
Matches 393; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 VLVDPPVDTGPKTYARDLPPKIDHYPGQILPTPKAIDFTPGQLGVSSELGLRPLYLES 66
DB 7 VLVDPPVDTGPKTYARDLPPKIDHYPGQILPTPKAIDFTPGQLGVSSELGLRPLYLES 66
QY 67 NGHTLVVTSKDGSDVFERELVDADVVISQPPWPAYLTPERTAKAKNKLATAGIGSD 126
DB 67 NGHTLVVTSKDGSDVFERELVDADVVISQPPWPAYLTPERTAKAKNKLATAGIGSD 126
QY 127 HVDLQSAIDBNVVAEYTCNSISVAEHVVMILSLVRNLPSEHWAARKGWNIAOCVSH 186
DB 127 HVDLQSAIDBNVVAEYTCNSISVAEHVVMILSLVRNLPSEHWAARKGWNIAOCVSH 186
QY 187 AYDLEAMHVGTVAAGRIGLAVLRRLAPFDVHLHYTDRHRLPESVEKELNTWHTREDMY 246
DB 187 AYDLEAMHVGTVAAGRIGLAVLRRLAPFDVHLHYTDRHRLPESVEKELNTWHTREDMY 246
QY 247 PVCDVYTLNCPLEPTEHMINDETLLKFRGAYIVNTARGKLCDRDAVARALESGRLAGY 306
DB 247 PVCDVYTLNCPLEPTEHMINDETLLKFRGAYIVNTARGKLCDRDAVARALESGRLAGY 306

Db 247 PVCDVYTLNCPLEPTEHMINDETLLKFRGAYIVNTARGKLCDRDAVARALESGRLAGY 306
QY 307 AGDVWFPQAPKDPHPTMTMYNGMTPHISGTTLTAQARVAGTRETILECFEGGRPIRDEY 366
Db 307 AGDVWFPQAPKDPHPTMTMYNGMTPHISGTTLTAQARVAGTRETILECFEGGRPIRDEY 366
QY 367 LIVOGLAAGTGAHSGYSKGNATGSGSEAAKFKKAV 401
Db 367 LIVOGLAAGTGAHSGYSKGNATGSGSEAAKFKKAV 401

Search completed: July 24, 2003, 20:59:04
Job time : 85 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:44:28 ; Search time 30 Seconds

(Without alignments)
565,555 Million cell updates/sec

Title: US-09-996-008B-2

Perfect score: 401
Sequence: 1 MAKVICLYDDPDVGYPKTV.....YSKGNATGSGSEAKFKKAV 401

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- Issued_Patents_AA:*
- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
 - 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
 - 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
 - 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
 - 5: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep:*
 - 6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	4.0	364	US-08-817-926-51	Sequence 51, Appl
2	16	4.0	364	US-09-203-893A-2	Sequence 2, Appl
3	16	4.0	364	US-09-203-893A-4	Sequence 4, Appl
4	16	4.0	364	US-09-203-893A-6	Sequence 6, Appl
5	16	4.0	364	US-09-203-893A-8	Sequence 8, Appl
6	16	4.0	364	US-09-203-893A-10	Sequence 10, Appl
7	16	4.0	364	US-09-203-893A-12	Sequence 12, Appl
8	16	4.0	364	US-09-203-893A-14	Sequence 14, Appl
9	16	4.0	364	US-09-203-893A-16	Sequence 16, Appl
10	16	4.0	364	US-09-203-893A-18	Sequence 18, Appl
11	16	4.0	364	US-09-203-893A-20	Sequence 20, Appl
12	16	4.0	364	US-09-203-893A-22	Sequence 22, Appl
13	16	4.0	364	US-09-203-893A-24	Sequence 24, Appl
14	16	4.0	364	US-09-203-893A-26	Sequence 26, Appl
15	16	4.0	364	US-09-203-893A-28	Sequence 28, Appl
16	16	4.0	364	US-09-203-893A-30	Sequence 30, Appl
17	16	4.0	364	US-09-203-893A-32	Sequence 32, Appl
18	16	4.0	253	US-09-530-058-6	Sequence 32, Appl
19	16	2.0	274	US-09-252-991A-24164	Sequence 2, Appl
20	8	2.0	550	US-08-844-058-2	Sequence 2, Appl
21	7	1.7	20	US-08-290-736C-16	Sequence 16, Appl
22	7	1.7	20	US-08-290-736C-27	Sequence 27, Appl
23	7	1.7	133	US-09-384-162-16	Sequence 16, Appl
24	7	1.7	155	US-09-252-991A-17465	Sequence 17, Appl
25	7	1.7	157	US-09-252-991A-26320	Sequence 26, Appl
26	7	1.7	176	US-09-462-842-3	Sequence 3, Appl
27	7	1.7	182	US-09-040-229B-4	Sequence 4, Appl

28	7	1.7	245	US-09-205-258-369	Sequence 369, App
29	7	1.7	259	US-09-252-991A-26709	Sequence 26709, A
30	7	1.7	268	US-08-965-056-71	Sequence 71, Appl
31	7	1.7	274	US-09-328-352-5762	Sequence 5762, Ap
32	7	1.7	304	US-09-634-238-238	Sequence 238, App
33	7	1.7	316	US-08-414-926A-22	Sequence 22, Appl
34	7	1.7	316	US-08-926-922-22	Sequence 22, Appl
35	7	1.7	316	US-09-253-682-22	Sequence 22, Appl
36	7	1.7	316	US-09-527-657-22	Sequence 22, Appl
37	7	1.7	331	US-09-252-991A-24420	Sequence 24420, A
38	7	1.7	348	US-09-107-532A-5765	Sequence 5765, Ap
39	7	1.7	365	US-09-252-991A-31439	Sequence 31439, A
40	7	1.7	366	US-09-252-991A-32332	Sequence 32332, A
41	7	1.7	394	US-08-867-030B-11	Sequence 11, Appl
42	7	1.7	394	PCT-US95-06119-11	Sequence 11, Appl
43	7	1.7	406	US-09-252-991A-19182	Sequence 19182, A
44	7	1.7	415	US-09-252-991A-19206	Sequence 19206, A
45	7	1.7	424	US-09-328-352-4427	Sequence 4427, Ap

ALIGNMENTS

RESULT 1
US-08-817-926-51
Sequence 51, Application US/08817926
Patent No. 6001590
GENERAL INFORMATION:
APPLICANT: Komeda, Yoshihiro
APPLICANT: Suda, Hisako
APPLICANT: Tamai, Yukio
APPLICANT: Iwamatsu, Akihito
APPLICANT: Kato, No. 6001590uo
APPLICANT: Sakai, Yasuyoshi
TITLE OF INVENTION: PROMOTER/TERMINATOR FOR CANDIDA BOLDINI
TITLE OF INVENTION: FORMATE DEHYDROGENASE GENE
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,926
FILING DATE: 09-MAY-1997
CLASSIFICATION: A35
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/02597
FILING DATE: 12-SEP-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 234133/1995
FILING DATE: 12-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 42536/1996
FILING DATE: 29-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 081356/0112
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 364 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-817-926-51

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVWFPQAPKDHPR 323
DB 281 GDVWFPQAPKDHPR 296

RESULT 2
US-09-203-893A-2
; Sequence 2, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of formate dehydrogenase from Candida
; TITLE OF INVENTION: boidin1, new Gene sequences encoding these and use of
; TITLE OF INVENTION: the new Formate dehydrogenases
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; PRIOR FILING DATE: 1998-12-02
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boidin1
US-09-203-893A-2

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVWFPQAPKDHPR 323
DB 281 GDVWFPQAPKDHPR 296

RESULT 3
US-09-203-893A-4
; Sequence 4, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of formate dehydrogenase from Candida
; TITLE OF INVENTION: boidin1, new Gene sequences encoding these and use of
; TITLE OF INVENTION: the new Formate dehydrogenases
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; PRIOR FILING DATE: 1998-12-02
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boidin1
US-09-203-893A-4

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVWFPQAPKDHPR 323
DB 281 GDVWFPQAPKDHPR 296

DB 281 GDVWFPQAPKDHPR 296

RESULT 4
US-09-203-893A-6
; Sequence 6, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of formate dehydrogenase from Candida
; TITLE OF INVENTION: boidin1, new Gene sequences encoding these and use of
; TITLE OF INVENTION: the new Formate dehydrogenases
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; PRIOR FILING DATE: 1998-12-02
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boidin1
US-09-203-893A-6

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVWFPQAPKDHPR 323
DB 281 GDVWFPQAPKDHPR 296

RESULT 5
US-09-203-893A-8
; Sequence 8, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of formate dehydrogenase from Candida
; TITLE OF INVENTION: boidin1, new Gene sequences encoding these and use of
; TITLE OF INVENTION: the new Formate dehydrogenases
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; PRIOR FILING DATE: 1998-12-02
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boidin1
US-09-203-893A-8

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVWFPQAPKDHPR 323
DB 281 GDVWFPQAPKDHPR 296

RESULT 6
US-09-203-893A-10
; Sequence 10, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of formate dehydrogenase from Candida
; TITLE OF INVENTION: boidin1, new Gene sequences encoding these and use of


```

; TITLE OF INVENTION: the new Formate dehydrogenases
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; CURRENT FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: DE 19753350.7
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boidinii
US-09-203-893A-10

Query Match
Best Local Similarity 100.0%; Score 16; DB 3; Length 364;
Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVWFPQAPKDPHWR 323
DB 281 GDVWFPQAPKDPHWR 296

RESULT 7
US-09-203-893A-12
; Sequence 12, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boidinii, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; CURRENT FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: DE 19753350.7
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boidinii
US-09-203-893A-12

Query Match
Best Local Similarity 100.0%; Score 16; DB 3; Length 364;
Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVWFPQAPKDPHWR 323
DB 281 GDVWFPQAPKDPHWR 296

RESULT 8
US-09-203-893A-14
; Sequence 14, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boidinii, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; CURRENT FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: DE 19753350.7
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 364
; TYPE: PRT
```

```

; ORGANISM: Candida boidinii
US-09-203-893A-14

Query Match
Best Local Similarity 100.0%; Score 16; DB 3; Length 364;
Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVWFPQAPKDPHWR 323
DB 281 GDVWFPQAPKDPHWR 296

RESULT 9
US-09-203-893A-16
; Sequence 16, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boidinii, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; CURRENT FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: DE 19753350.7
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boidinii
US-09-203-893A-16

Query Match
Best Local Similarity 100.0%; Score 16; DB 3; Length 364;
Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVWFPQAPKDPHWR 323
DB 281 GDVWFPQAPKDPHWR 296

RESULT 10
US-09-203-893A-18
; Sequence 18, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boidinii, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; CURRENT FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: DE 19753350.7
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boidinii
US-09-203-893A-18

Query Match
Best Local Similarity 100.0%; Score 16; DB 3; Length 364;
Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVWFPQAPKDPHWR 323
DB 281 GDVWFPQAPKDPHWR 296
```

RESULT 11
US-09-203-893A-20
; Sequence 20, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boldin1, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; PRIOR FILING DATE: 1998-12-02
; PRIOR FILING DATE: 19753350.7
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 20
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boldin1
US-09-203-893A-20

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVFPQAPKDPHWR 323
Db 281 GDVFPQAPKDPHWR 296

RESULT 12
US-09-203-893A-22
; Sequence 22, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boldin1, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; PRIOR FILING DATE: 1998-12-02
; PRIOR FILING DATE: 19753350.7
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 22
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boldin1
US-09-203-893A-22

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVFPQAPKDPHWR 323
Db 281 GDVFPQAPKDPHWR 296

RESULT 13
US-09-203-893A-24
; Sequence 24, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boldin1, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012

; CURRENT APPLICATION NUMBER: US/09/203,893A
; CURRENT FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: DE 19753350.7
; PRIOR FILING DATE: 1997-12-03
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 24
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boldin1
US-09-203-893A-24

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVFPQAPKDPHWR 323
Db 281 GDVFPQAPKDPHWR 296

RESULT 14
US-09-203-893A-26
; Sequence 26, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boldin1, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; PRIOR FILING DATE: 1998-12-02
; PRIOR FILING DATE: 19753350.7
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 26
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boldin1
US-09-203-893A-26

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVFPQAPKDPHWR 323
Db 281 GDVFPQAPKDPHWR 296

RESULT 15
US-09-203-893A-28
; Sequence 28, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boldin1, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; PRIOR FILING DATE: 1998-12-02
; PRIOR FILING DATE: 19753350.7
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 28
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boldin1
US-09-203-893A-28

Query Match 4.0%; Score 16; DB 3; Length 364;
 Best Local Similarity 100.0%; Pred. No. 2.1e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 308 GDVWFPOPAPKDHPR 323
 ||||||||||||
 Db 281 GDVWFPOPAPKDHPR 296

Search completed: July 24, 2003, 20:49:56
 Job time : 42 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:48:38 ; Search time 54 seconds
(without alignments)
881.901 Million cell updates/sec

Title: US-09-996-008B-2

Perfect score: 401

Sequence: 1 MAKVLCVLYDDPVDGYPPTY.....YSGNATGSGSEAKFKKAV 401

Scoring table:

GAPOP 60.0 , Gapext 60.0

Searched: 451899 seqs, 118759770 residues

Word size : 0

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PC1_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PC1US_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	4.7	387	US-10-156-761-9479	Sequence 9479, Ap
2	11	2.7	376	US-10-224-567-1	Sequence 1, Appli
3	9	2.2	333	US-10-156-761-12108	Sequence 12108, A
4	8	2.0	335	US-09-815-242-5830	Sequence 5830, Ap
5	8	2.0	343	US-09-815-242-12981	Sequence 12981, A
6	8	2.0	398	US-09-815-242-13150	Sequence 13150, A
7	8	2.0	617	US-09-712-363-160	Sequence 160, App
8	8	2.0	617	US-09-815-242-13228	Sequence 13228, A
9	8	2.0	764	US-10-066-551-2	Sequence 2, Appli
10	8	2.0	1176	US-10-156-761-11086	Sequence 11086, A
11	7	1.7	38	US-10-099-766-3	Sequence 3, Appli
12	7	1.7	40	US-09-682-706-6	Sequence 6, Appli
13	7	1.7	161	US-09-682-706-5	Sequence 5, Appli
14	7	1.7	161	US-10-099-766-2	Sequence 2, Appli
15	7	1.7	162	US-10-156-761-13701	Sequence 13701, A

16	7	1.7	175	US-10-156-761-14883	Sequence 14883, A
17	7	1.7	176	US-10-339-278-3	Sequence 3, Appli
18	7	1.7	182	US-10-156-749-4	Sequence 4, Appli
19	7	1.7	218	US-10-102-806-515	Sequence 515, App
20	7	1.7	220	US-09-947-316-2	Sequence 2, Appli
21	7	1.7	231	US-09-934-455-286	Sequence 286, App
22	7	1.7	245	US-09-799-777-57	Sequence 57, Appli
23	7	1.7	245	US-09-947-316-1	Sequence 1, Appli
24	7	1.7	245	US-10-022-282-369	Sequence 369, App
25	7	1.7	258	US-10-156-761-7639	Sequence 7639, Ap
26	7	1.7	268	US-09-854-816-71	Sequence 71, Appli
27	7	1.7	268	US-10-156-761-15022	Sequence 15022, A
28	7	1.7	301	US-09-971-536-51	Sequence 51, Appli
29	7	1.7	363	US-10-156-761-13930	Sequence 13930, A
30	7	1.7	478	US-10-156-761-13018	Sequence 13018, A
31	7	1.7	511	US-09-815-242-5147	Sequence 5147, Ap
32	7	1.7	513	US-09-833-745-64	Sequence 64, Appli
33	7	1.7	637	US-09-817-310-2	Sequence 2, Appli
34	7	1.7	658	US-09-815-242-10947	Sequence 10947, A
35	7	1.7	663	US-10-156-761-15035	Sequence 15035, A
36	7	1.7	679	US-09-728-626-5995	Sequence 5995, Ap
37	7	1.7	692	US-10-156-761-14649	Sequence 14649, A
38	7	1.7	731	US-10-205-823-202	Sequence 202, Appli
39	7	1.7	758	US-09-735-101-2	Sequence 5, Appli
40	7	1.7	1203	US-09-799-875-5	Sequence 2, Appli
41	7	1.7	1257	US-10-109-324-2	Sequence 3707, Ap
42	7	1.7	1510	US-09-728-626-3707	Sequence 13716, A
43	7	1.7	1516	US-10-156-761-13716	Sequence 43, Appli
44	7	1.7	1724	US-09-964-899-43	Sequence 12, Appli
45	7	1.7	1925	US-10-205-032-12	

ALIGNMENTS

RESULT 1
US-10-156-761-9479
Sequence 9479, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 9479
LENGTH: 387
TYPE: PR1
ORGANISM: Streptomyces avermitilis
US-10-156-761-9479

Query Match 4.7%; Score 19; DB 15; Length 387;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 IDFTGQLGSGVSGELGLR 61
|||||
Db 43 IDFTGQLGSGVSGELGLR 61

RESULT 2
US-10-224-567-1

```
; Sequence 1, Application US/10224567
; Publication No. US20030087763A1
; GENERAL INFORMATION:
; APPLICANT: Kobayashi, Akio
; APPLICANT: Fukusaki, Eiichiro
; APPLICANT: Isogai, Akira
; TITLE OF INVENTION: Method to Promote Growth of a Plant
; FILE REFERENCE: 026350-077
; CURRENT APPLICATION NUMBER: US/10/224,567
; CURRENT FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: US 09/517,427
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: JP 11-56,776
; PRIOR FILING DATE: 1999-03-04
; PRIOR APPLICATION NUMBER: JP 2000-36,153
; PRIOR FILING DATE: 2000-02-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Rice formate dehydrogenase
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 287
; OTHER INFORMATION: Xaa - Any Amino Acid
US-10-224-567-1
```

```
Query Match 2.7%; Score 11; DB 15; Length 376;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 308 GDVWFPQAPK 318
Db 305 GDVWFPQAPK 315
```

```
RESULT 3
US-10-156-761-12108
; Sequence 12108, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMODA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12108
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12108
```

```
Query Match 2.2%; Score 9; DB 15; Length 333;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 290 DRDAVARAL 298
Db 42 DRDAVARAL 50
```

```
RESULT 4
US-09-815-242-5830
; Sequence 5830, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5830
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5830
```

```
Query Match 2.0%; Score 8; DB 9; Length 335;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 319 DHPRTMP 326
Db 277 DHPRTMP 284
```

```
RESULT 5
US-09-815-242-12981
; Sequence 12981, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
```

PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12981
LENGTH: 343
TYPE: PRT
ORGANISM: *Staphylococcus aureus*
US-09-815-242-12981

Query Match 2.0%; Score 8; DB 9; Length 343;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 DHPWRTMP 326
DB 280 DHPWRTMP 287

RESULT 6
US-09-815-242-13150
Sequence 13150, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlson, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13150
LENGTH: 343
TYPE: PRT
ORGANISM: *Staphylococcus aureus*
US-09-815-242-13150

Query Match 2.0%; Score 8; DB 9; Length 343;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 DHPWRTMP 326
DB 280 DHPWRTMP 287

RESULT 7
US-09-712-363-160
Sequence 160, Application US/09712363
Patent No. US20020164588A1
GENERAL INFORMATION:
APPLICANT: Eisenberg, David
APPLICANT: Marcotte, Edward M.
APPLICANT: Rotstein, Sergio H.
TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
FILE REFERENCE: 07419-032001
CURRENT APPLICATION NUMBER: US/09/712,363
PRIOR FILING DATE: 2000-11-13
PRIOR APPLICATION NUMBER: PCT/US00/02246
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: 60/179,531
PRIOR FILING DATE: 2000-02-01
PRIOR APPLICATION NUMBER: 60/117,844
PRIOR FILING DATE: 1999-01-29
PRIOR APPLICATION NUMBER: 60/118,206,
PRIOR FILING DATE: 1999-02-01
PRIOR APPLICATION NUMBER: 60/126,593
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/134,093
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/134,092
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/165,124
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/165,086
NUMBER OF SEQ ID NOS: 292
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 160
LENGTH: 398
TYPE: PRT
ORGANISM: *Mycobacterium tuberculosis*
US-09-712-363-160

Query Match 2.0%; Score 8; DB 10; Length 398;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 AVLRRLAP 213
DB 181 AVLRRLAP 188

RESULT 8
US-09-815-242-13228
Sequence 13228, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlson, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26

```
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 1410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13228
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13228
```

```
Query Match          2.0%; Score 8; DB 9; Length 617;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      211 LAPDVHL 218
        |||||||
DB      518 LAPDVHL 525
```

```
RESULT 9
US-10-066-551-2
; Sequence 2, Application US/10066551
; Publication No. US20030100071A1
; GENERAL INFORMATION:
; APPLICANT: Apicella, M. A.
; APPLICANT: Edwards, J. L.
; APPLICANT: Gibson, B. W.
; APPLICANT: Scheffler, K.
; APPLICANT: Brown, E.
; TITLE OF INVENTION: Vaccine and compositions for the
; TITLE OF INVENTION: Prevention and treatment of Neisserial infections
; FILE REFERENCE: 875.045051
; CURRENT APPLICATION NUMBER: US/10/066,551
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/344,452
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/310,356
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/266,070
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 12
; SEQ ID NO 2
; LENGTH: 764
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-066-551-2
```

```
Query Match          2.0%; Score 8; DB 15; Length 764;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      54 VSGELGR 61
        |||||||
DB      178 VSGELGR 185
```

```
RESULT 10
US-10-156-761-11086
; Sequence 11086, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
```

```
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11086
; LENGTH: 1176
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11086
```

```
Query Match          2.0%; Score 8; DB 15; Length 1176;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      205 LAVLRRLA 212
        |||||||
DB      106 LAVLRRLA 113
```

```
RESULT 11
US-10-099-766-3
; Sequence 3, Application US/10099766
; Publication No. US20030036507A1
; GENERAL INFORMATION:
; APPLICANT: Lewis, Kathy
; APPLICANT: Vale, Wylie
; APPLICANT: Marilyn H. Perrin
; APPLICANT: Jean E. Rivier
; APPLICANT: Koichi S. Kunitake
; APPLICANT: Joszef Gulyas
; TITLE OF INVENTION: Urocortin III and uses thereof
; FILE REFERENCE: D6390
; CURRENT APPLICATION NUMBER: US/10/099,766
; CURRENT FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/276,069
; PRIOR FILING DATE: 2001-03-15
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 3
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: Human urocortin III (hucn III)
US-10-099-766-3
```

```
Query Match          1.7%; Score 7; DB 15; Length 38;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      109 IAKAKNL 115
        |||||||
DB      18 IAKAKNL 24
```

```
RESULT 12
US-09-682-706-6
; Sequence 6, Application US/09682706
; Patent No. US20020082409A1
; GENERAL INFORMATION:
; APPLICANT: Hsu, Sheau-Yu
; APPLICANT: Hsueh, Aaron
; TITLE OF INVENTION: Stresscoplins and their ses
; FILE REFERENCE: STAN210
; CURRENT APPLICATION NUMBER: US/09/682,706
; CURRENT FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/276,615
; PRIOR FILING DATE: 2001-03-15
```


PRIOR APPLICATION NUMBER: 60/244,128
PRIOR FILING DATE: 2000-10-26
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 40
TYPE: PRT
ORGANISM: Homo sapiens
US-09-682-706-6

Query Match 1.7%; Score 7; DB 9; Length 40;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 IAKAKNL 115
Db 20 IAKAKNL 26

RESULT 13
US-09-682-706-5
Sequence 5, Application US/09682706
Patent No. US20020082409A1
GENERAL INFORMATION:
APPLICANT: Hsu, Sheau-Yu
APPLICANT: Hsueh, Aaron
TITLE OF INVENTION: Stresscopins and their ses
FILE REFERENCE: STRAN210
CURRENT APPLICATION NUMBER: US/09/682,706
CURRENT FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/276,615
PRIOR FILING DATE: 2001-03-15
PRIOR APPLICATION NUMBER: 60/244,128
PRIOR FILING DATE: 2000-10-26
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 161
TYPE: PRT
ORGANISM: Homo sapiens
US-09-682-706-5

Query Match 1.7%; Score 7; DB 9; Length 161;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 IAKAKNL 115
Db 137 IAKAKNL 143

RESULT 14
US-10-099-766-2
Sequence 2, Application US/10099766
Publication No. US20030036507A1
GENERAL INFORMATION:
APPLICANT: Lewis, Kathy
APPLICANT: Vale, Wylie
APPLICANT: Marilyn H. Perrin
APPLICANT: Jean E. Rivier
APPLICANT: Koichi S. Kunitake
APPLICANT: Jozsef Gulyas
TITLE OF INVENTION: Urocortin III and Uses thereof
FILE REFERENCE: D6390
CURRENT APPLICATION NUMBER: US/10/099,766
CURRENT FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: US 60/276,069
PRIOR FILING DATE: 2001-03-15
NUMBER OF SEQ ID NOS: 17
SEQ ID NO 2
LENGTH: 161
TYPE: PRT
ORGANISM: Homo sapiens

FEATURE:
OTHER INFORMATION: Human urocortin III Precursor
US-10-099-766-2

Query Match 1.7%; Score 7; DB 15; Length 161;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 IAKAKNL 115
Db 137 IAKAKNL 143

RESULT 15
US-10-156-761-13701
Sequence 13701, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 13701
LENGTH: 162
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-13701

Query Match 1.7%; Score 7; DB 15; Length 162;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 AVLRLA 212
Db 32 AVLRLA 38

Search completed: July 24, 2003, 20:57:39
Job time : 62 secs

A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, hebaullt, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: Sma0478
A;Genome: plasmid
C;Superfamily: Neurospora formate dehydrogenase

Query Match 6.0%; Score 24; DB 2; Length 401;
Best Local Similarity 100.0%; Pred. No. 7.2e-16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 GHTLVTSKDGSDSVFERELVDA 91
|||||
Db 70 GHTLVTSKDGSDSVFERELVDA 93

RESULT 3

JC7815
formate dehydrogenase (EC 1.2.1.2) - *Paracoccus* sp. (Strain 12-A)
C;Species: *Paracoccus* sp.
C;Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 03-Jun-2002
C;Accession: JC7815
R;Shinoda, T.; Satoh, T.; Mineki, S.; Iida, M.; Taguchi, H. Biosci. Biotechnol. Biochem. 66, 271-276, 2002
A;Title: Cloning, nucleotide sequencing, and expression in *Escherichia coli* of the gene
A;Reference number: JC7815; PMID:11999398; MUID:21994041
A;Accession: JC7815
A;Molecule type: DNA
A;Residues: 1-400 <SH1>
A;Cross-references: DDBJ:AB071373
C;Comment: This enzyme, which catalyzes the conversion of formate to carbon dioxide with ti-enzyme systems, such as bioreactors involving NADH as a coenzyme, and plays a key rol
C;Genetics:
A;Gene: fdh
C;Keywords: oxidoreductase

Query Match 5.7%; Score 23; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 7.4e-15;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 378 GAHSYSGNATGGSEEAFFKA 400
|||||
Db 378 GAHSYSGNATGGSEEAFFKA 400

RESULT 4

JC4252
formate dehydrogenase (EC 1.2.1.2) - yeast (*Candida methyllica*)
C;Species: *Candida methyllica*
C;Date: 19-Oct-1995 #sequence_revision 08-Feb-1996 #text_change 23-Sep-2002
C;Accession: JC4252; PC4070
R;Allen, S.J.; Holbrook, J.J. Gene 162, 99-104, 1995

A;Title: Isolation, sequence and overexpression of the gene encoding NAD-dependent formate
A;Reference number: JC4252; MUID:96009885; PMID:7557425
A;Accession: JC4252

A;Molecule type: DNA
A;Residues: 1-364 <ALL>
A;Cross-references: EMBL:X81129; NID:g1181203; PIDN:CAA57036.1; PID:g1181204
A;Accession: PC4070

A;Molecule type: protein
A;Residues: 1-30 <AL2>
A;Experimental source: ATCC 56294
A;Comment: This enzyme catalyses the final step in the dissimilatory pathway of methanol of NAD+ to NADH.
C;Genetics:
A;Gene: fdh

C;Superfamily: Neurospora formate dehydrogenase
C;Keywords: homodimer; NAD; oxidoreductase
F;165-195/Region: beta-alpha-beta NAD nucleotide-binding fold

Query Match 4.0%; Score 16; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 8.4e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVWFPPQAPKDPWR 323
|||||
Db 281 GDVWFPPQAPKDPWR 296

RESULT 5

S65308
hypothetical protein YPL275w - yeast (*Saccharomyces cerevisiae*)
N;Alternate names: hypothetical protein P0326
C;Species: *Saccharomyces cerevisiae*
C;Date: 10-Dec-1994 #sequence_revision 31-May-1996 #text_change 23-Sep-2002
C;Accession: S65308; S65329
R;Duesterhoeft, A.; Floeth, M.; Fritze, M.; Hilbert, H.; Moestl, D. submitted to the Protein Sequence Database, May 1996
A;Reference number: S65292
A;Accession: S65308
A;Molecule type: DNA
A;Residues: 1-236 <DUE>
A;Cross-references: EMBL:273632; NID:g1370566; PIDN:CAA98013.1; PID:g1370568; MIPS:YP
A;Experimental source: strain S288C (AB972)
R;Dellius, H.; Hebling, U. submitted to the Protein Sequence Database, May 1996
A;Reference number: S64967
A;Accession: S65329
A;Molecule type: DNA
A;Residues: 1-236
A;Cross-references: EMBL:273632; NID:g1370566; PIDN:CAA98013.1; PID:g1370568; MIPS:YP
A;Experimental source: strain S288C (AB972)
C;Genetics:
A;Cross-references: SGD:S0006196
A;Map position: 16L
C;Superfamily: Neurospora formate dehydrogenase

Query Match 3.0%; Score 12; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 314 QPAPKDPWRMT 325
|||||
Db 161 QPAPKDPWRMT 172

RESULT 6

S67300
probable membrane protein YOR388c - yeast (*Saccharomyces cerevisiae*)
N;Alternate names: hypothetical protein O6778
C;Species: *Saccharomyces cerevisiae*
C;Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 23-Sep-2002
C;Accession: S67300
R;Dellius, H.; Hebling, U.; Hofmann, B. submitted to the Protein Sequence Database, July 1996
A;Reference number: S67261
A;Accession: S67300

A;Molecule type: DNA
A;Residues: 1-376
A;Cross-references: EMBL:275296; NID:g1420834; PIDN:CAA99720.1; PID:g1420835; MIPS:YO
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:FDH1
A;Cross-references: SGD:S0005915; MIPS:YOR388c
A;Map position: 15R

C;Superfamily: Neurospora formate dehydrogenase
C;Keywords: transmembrane protein
F;121-137/Domain: transmembrane #status predicted <TM>

Query Match 3.0%; Score 12; DB 2; Length 376;
Best Local Similarity 100.0%; Pred. No. 0.00098;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 314 QPAPKDPHPRM 325
 DB 301 QPAPKDPHPRM 312

RESULT 7

formate dehydrogenase (EC 1.2.1.2) precursor, mitochondrial - potato
 C:Species: Solanum tuberosum (potato)
 C>Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 23-Sep-2002
 C:Accession: J02272; PC2189; S3125
 R:des France-Small, C.C.; Ambard-Bretteville, F.; Small, I.D.; Remy, R.
 plant physiol 102, 1171-1177, 1993
 A>Title: Identification of a major soluble protein in mitochondria from nonphotosynthetic
 A:Reference number: J02272; MUID:94105343; PMID:8278546
 A:Accession: J02272
 A:Molecule type: mRNA
 A:Residues: 1-379 <DESI>
 A:Cross-references: EMBL:Z21493; NID:g297797; PIDN:CAAT9702.1; PID:g297798
 A:Accession: PC2189
 A:Molecule type: protein
 A:Residues: 24-50 <DESI>
 C:Superfamily: Neurospora formate dehydrogenase
 C:Keywords: mitochondrion; oxidoreductase
 F:1-23/Domain: transit peptide (mitochondrion) #status predicted <TM>
 F:24-379/Product: formate dehydrogenase #status predicted <MAT>
 F:199-222/Region: NAD binding #status predicted

Query Match

Best Local Similarity 3.0%; Score 12; DB 2; Length 379;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 119 LTAGISDHYDL 130
 DB 119 LTAGISDHYDL 130

RESULT 8

formate dehydrogenase (EC 1.2.1.2) - Neurospora crassa
 C:Species: Neurospora crassa
 C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 23-Sep-2002
 C:Accession: A47117
 R:Chow, C.M.; Rajbhandary, U.L.
 J. Bacteriol. 175, 3703-3709, 1993
 A>Title: Developmental regulation of the gene for formate dehydrogenase in Neurospora cr
 A:Reference number: A47117; MUID:93285982; PMID:8509325
 A:Accession: A47117
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-375 <CHO>
 A:Cross-references: GB:L13964; NID:g293949; PIDN:AAA99900.1; PID:g13121604
 C:Genetics:
 A:Introns: 1/3; 17/3
 C:Superfamily: Neurospora formate dehydrogenase
 C:Keywords: oxidoreductase

Query Match 2.7%; Score 11; DB 2; Length 375;
 Best Local Similarity 100.0%; Pred. No. 0.01;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 120 TAGISDHYDL 130
 DB 91 TAGISDHYDL 101

RESULT 9

S30088
 C:Species: Emerizella nidulans
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 23-Sep-2002
 C:Accession: S30088

R:Saleeba, J.A.; Cobbett, C.S.; Hynes, M.J.

Mol. Gen. Genet. 235, 349-358, 1992

A>Title: Characterization of the amA-regulated acIA gene of Aspergillus nidulans.

A:Reference number: S30088; MUID:93101140; PMID:1465107

A:Accession: S30088

A:Molecule type: DNA

A:Residues: 1-377 <SAT>

A:Cross-references: EMBL:J11612; NID:g5544; PIDN:CAAT7687.1; PID:g5545

C:Genetics:

A:Gene: acIA

A:Map position: 1

A:Introns: 5/3; 64/2

C:Superfamily: Neurospora formate dehydrogenase

Query Match 2.7%; Score 11; DB 2; Length 377;
 Best Local Similarity 100.0%; Pred. No. 0.01;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 120 TAGISDHYDL 130
 DB 86 TAGISDHYDL 96

RESULT 10

formate dehydrogenase (FDH) - Arabidopsis thaliana
 T51423
 N:Alternate names: protein T9L3_80
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 23-Sep-2002
 C:Accession: T51423
 R:Saito, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.;
 submitted to the Protein Sequence Database, August 2000
 A:Reference number: Z25394
 A:Accession: T51423
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-384 <SAT>
 A:Cross-references: EMBL:AL391149
 A:Experimental source: cultivar Columbia; BAC clone T9L3
 C:Genetics:
 A:Map position: 5
 A:Introns: 29/3; 88/1; 268/2; 347/3
 A>Note: T9L3 80
 C:Superfamily: Neurospora formate dehydrogenase

Query Match 2.7%; Score 11; DB 2; Length 384;
 Best Local Similarity 100.0%; Pred. No. 0.01;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 313 QPAPKDPHPR 323
 DB 318 QPAPKDPHPR 328

RESULT 11

conserved hypothetical protein [imported] -- Sulfolobus solfataricus
 C:Species: Sulfolobus solfataricus
 C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
 C:Accession: D90271
 R:She, Q.; Sindh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Ayañez, M.J.; Ch
 Jong, I.; Jeffries, A.C.; Kozera, C.U.; Medina, N.; Peng, X.; Thl-Ngoc, H.P.; Redder
 arett, R.A.; Ragan, M.A.; Jensen, C.W.; Van der Oost, J.
 submitted to GenBank, April 2001
 A:Description: Sulfolobus solfataricus complete genome.
 A:Reference number: A99139
 A:Accession: D90271
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-89 <KUR>
 A:Cross-references: GB:AE006641; NID:g13814371; PIDN:AAK41427.1; GSPDB:GN00155
 C:Genetics:
 A:Gene: SSO7348

Query Match 2.0%; Score 8; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 197 TVAAGRIG 204
DB 19 TVAAGRIG 26

RESULT 12

hypothetical protein Atu0077 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C:Accession: AH2585
R:Wood, D.W.; Sebald, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayvin, T.; Levy, R.; Li, M.; McCell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AH2585
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-334 <KUR>
A:Cross-references: GB:AE008688; PIDN:AL41102.1; PID:q17738393; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu0077
A:Map position: circular chromosome
C:Superfamily: phosphoglycerate dehydrogenase

Query Match 2.0%; Score 8; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 YIVNTARG 286
DB 240 YIVNTARG 247

RESULT 13

G97367

probable dehydrogenase [imported] - Agrobacterium tumefaciens (strain C58, Cereon)

C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C:Accession: G97367
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quirillo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: G97367
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-334 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK85896.1; PID:q15154941; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_114
A:Map position: circular chromosome
C:Superfamily: phosphoglycerate dehydrogenase

Query Match 2.0%; Score 8; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 YIVNTARG 286
DB 240 YIVNTARG 247

RESULT 14
E89779
NAD-dependent formate dehydrogenase [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: E89779
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsugu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: E89779
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-374 <KUR>
A:Cross-references: GB:BA000018; PID:q13700093; PIDN:BA841392.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: fdh

Query Match 2.0%; Score 8; DB 2; Length 374;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 DHPWRTMP 326
DB 311 DHPWRTMP 318

RESULT 15

B70752

hypothetical protein RV0106 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: B70752
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: B70752
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-398 <COL>
A:Cross-references: GB:Z74410; GB:AL123456; NID:q3261600; PIDN:CA98919.1; PID:q14057
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV0106
C:Superfamily: conserved hypothetical protein yjc

Query Match 2.0%; Score 8; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 206 AVLRLAP 213
DB 181 AVLRLAP 188

Search completed: July 24, 2003, 21:02:09
Job time: 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:50:03 ; Search time 22 Seconds

(without alignments)
857.169 Million cell updates/sec

Title: US-09-996-008b-2
Sequence: 1 MAKVLCVLYDDPDVGYPKTY.....YKGNATGSEAKFKKAV 401

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size: 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: SwissProt_41*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	339	84.5	400	1	FDH_PSESER
2	16	4.0	361	1	FDH_PICAN
3	16	4.0	377	1	FDH_HORVU
4	12	3.0	381	1	FDH_SOLUT
5	11	2.7	375	1	FDH_NEUCR
6	11	2.7	376	1	FDH_ORISA
7	11	2.7	377	1	FDH_EMENT
8	11	2.7	384	1	FDH_ARATH
9	8	2.0	190	1	FRDA_DROME
10	8	2.0	398	1	Y106_MYCTU
11	8	2.0	430	1	PURA_XANAC
12	8	2.0	430	1	PURA_XANAC
13	8	2.0	430	1	PURA_XANAC
14	7	1.7	126	1	PDDB_SULSO
15	7	1.7	134	1	Y605_METVA
16	7	1.7	161	1	UCN3_HUMAN
17	7	1.7	176	1	DSBB_ECOL6
18	7	1.7	176	1	DSBB_ECOLI
19	7	1.7	176	1	DSBB_SHITL
20	7	1.7	206	1	COAE_HAETN
21	7	1.7	211	1	RL1_HALCU
22	7	1.7	277	1	MORI_HALNI
23	7	1.7	287	1	MORI_BRUME
24	7	1.7	293	1	RRSK_BACSU
25	7	1.7	293	1	RRSK_BACSU
26	7	1.7	309	1	DHL2_LACCO
27	7	1.7	314	1	PIX2_BRARE
28	7	1.7	314	1	RLAO_THEAC
29	7	1.7	329	1	YL31_ARCFU
30	7	1.7	334	1	YEAA_SCHPO
31	7	1.7	350	1	YNI4_YEAST
32	7	1.7	391	1	CAL1_YEAST
33	7	1.7	394	1	UDG_STRPN
33	7	1.7	398	1	HEM1_METTM

34	7	1.7	402	1	HEM1_METTM
35	7	1.7	402	1	OPS4_CANAL
36	7	1.7	403	1	KICS_MOUSE
37	7	1.7	408	1	PEPT_CLOAB
38	7	1.7	414	1	SVS2_RAT
39	7	1.7	418	1	HLT_VIBPA
40	7	1.7	427	1	PLT1_HUMAN
41	7	1.7	428	1	PLT1_MOUSE
42	7	1.7	428	1	PLT1_RAT
43	7	1.7	463	1	GUN_BACSP
44	7	1.7	464	1	KTR4_YEAST
45	7	1.7	486	1	MENE_BACSU

ALIGNMENTS

RESULT 1	FDH_PSESER	STANDARD:	PRT:	400 AA.
AC	P33160:			
DT	01-OCT-1993 (Rel. 27, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Formate dehydrogenase (EC 1.2.1.2) (NAD-dependent formate dehydrogenase) (FDH).			
DE	Pseudomonas sp. (strain 101) (Achromobacter parvulus T1).			
OC	Bacteria; Proteobacteria.			
OX	NCBI_TaxID=33067;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=92063805; PubMed=1954846;			
RA	Tishkov V.I., Galikin A.G., Egorov A.M.;			
RT	"NAD-dependent formate dehydrogenase of methylotrophic bacteria			
RT	Pseudomonas sp. 101: cloning, expression, and study of the genetic			
RT	structure.";			
RL	Dokl. Akad. Nauk SSSR 317:745-748(1991).			
RN	[2]			
RP	SEQUENCE OF 1-393.			
RX	MEDLINE=90290536; PubMed=2357236;			
RA	Popov V.O., Shumilin I.A., Ustinaikova T.B., Lamzin V.S., Egorov T.A.;			
RT	"NAD-dependent formate dehydrogenase from methylotrophic bacteria			
RT	Pseudomonas sp. 101.I. Amino acid sequence.";			
RL	Bioorg. Khim. 16:324-335(1990).			
RN	[3]			
RP	X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).			
RX	MEDLINE=92283270; PubMed=1597184;			
RA	Lamzin V.S., Aleshin A.E., Strokopytov B.V., Yukhnevlch M.G.,			
RA	Popov V.O., Harutyunyan E.H., Wilson K.S.;			
RT	"Crystal structure of NAD-dependent formate dehydrogenase.";			
RL	Eur. J. Biochem. 206:441-452(1992).			
RN	[4]			
RP	MUTAGENESIS OF CYS-255.			
RX	MEDLINE=93249485; PubMed=8484798;			
RA	Tishkov V.I., Galikin A.G., Marchenko G.N., Egorova O.A., Sheluho D.V.,			
RA	Kulakova L.B., Dementieva L.A., Egorov A.M.;			
RT	"Catalytic properties and stability of a Pseudomonas sp.101 formate			
RT	dehydrogenase mutants containing Cys-255-Ser and Cys-255-Met			
RT	replacements.";			
RL	Biochem. Biophys. Res. Commun. 192:976-981(1993).			
CC	"- CATALYTIC ACTIVITY: Formate + NAD(+) - CO(2) + NADH.			
CC	"- SUBUNIT: Homodimer			
CC	"- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID			
CC	DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.			
CC	PIR: J00334; J00334.			
DR	PDB: 2NAD; 26-JAN-95.			
DR	InterPro: IPR006139; 2-Hacid_DH.			
DR	InterPro: IPR006140; 2-Hacid_DH.C.			
DR	Plan: PF00389; 2-Hacid_DH; 1.			
DR	Plan: PF02826; 2-Hacid_DH.C; 1.			
DR	PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.			
DR	PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.			

DR PROSITE; PS00671; D_2-HYDROXYACID_DH_3; 1.
 KW Oxidoreductase; NAD; 3D-structure.
 FT INIT_MET 0 0
 FT DOMAIN 1 146 CATALYTIC.
 FT DOMAIN 147 333 COENZYME-BINDING.
 FT DOMAIN 334 400 CATALYTIC.
 FT NP_BIND 192 226 NAD (ADP PART).
 FT ACT_SITE 5 5 POTENTIAL.
 FT ACT_SITE 284 284 SUBSTRATE BINDING.
 FT MUTAGEN 255 255 C->S,M: HIGH RESISTANCE TO INACTIVATION
 BY HG(2+); INCREASED STABILITY AT 25 C
 AND DECREASED THERMOSTABILITY AT 45 C.
 D->S (IN REF. 1).
 TV->VT (IN REF. 1).
 C->V (IN REF. 1).
 VH->HV (IN REF. 1).
 N->D (IN REF. 1).
 FT STRAND 2 6
 FT STRAND 8 8
 FT TURN 12 13
 FT TURN 31 32
 FT TURN 46 47
 FT STRAND 52 52
 FT TURN 53 55
 FT TURN 56 58
 FT HELIX 60 65
 FT TURN 66 67
 FT STRAND 69 73
 FT TURN 79 80
 FT HELIX 82 87
 FT TURN 88 89
 FT STRAND 92 96
 FT TURN 97 98
 FT STRAND 99 99
 FT STRAND 103 103
 FT HELIX 105 110
 FT TURN 112 113
 FT STRAND 116 119
 FT TURN 125 126
 FT STRAND 127 127
 FT HELIX 129 134
 FT TURN 135 136
 FT STRAND 138 141
 FT TURN 143 144
 FT HELIX 147 162
 FT TURN 163 164
 FT HELIX 165 173
 FT TURN 174 175
 FT HELIX 179 183
 FT TURN 184 185
 FT TURN 190 191
 FT STRAND 193 197
 FT HELIX 201 210
 FT TURN 211 213
 FT STRAND 214 214
 FT STRAND 216 220
 FT HELIX 227 233
 FT TURN 234 234
 FT STRAND 235 237
 FT HELIX 241 244
 FT STRAND 245 247
 FT STRAND 250 253
 FT TURN 259 263
 FT STRAND 265 265
 FT HELIX 267 270
 FT TURN 271 272
 FT TURN 275 276
 FT STRAND 277 281
 FT HELIX 285 287
 FT STRAND 288 288
 FT HELIX 290 298
 FT TURN 299 300
 FT STRAND 302 307

FT TURN 317 318
 FT HELIX 320 322
 FT TURN 323 323
 FT TURN 325 326
 FT STRAND 327 327
 FT HELIX 334 336
 FT HELIX 338 357
 FT TURN 358 358
 FT HELIX 363 365
 FT STRAND 366 368
 FT STRAND 373 373
 FT TURN 375 375
 FT HELIX 376 380
 SQ SEQUENCE 400 AA; 44005 MW; A354151D4312991F CRC64;
 Query Match 84.5%; Score 339; DB 1; Length 400;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 63 YLESGHTLVVTSKDGSPSFERELVDADVVISQFPVAYLTPERIARAKNLKALTAG 122
 DB 62 YLESGHTLVVTSKDGSPSFERELVDADVVISQFPVAYLTPERIARAKNLKALTAG 121
 QY 123 IGSDDVDLOSADIRNTVAEVTYCNLSVAEHVMMILSLVNNYLPSEHARKGNMID 182
 DB 122 IGSDDVDLOSADIRNTVAEVTYCNLSVAEHVMMILSLVNNYLPSEHARKGNMID 181
 QY 183 CVSHAYDLEAMHYGTVAAAGTIGLAVLRRLAPFDVHLHYDHRHLPSEVEKELNLTWHATR 242
 DB 182 CVSHAYDLEAMHYGTVAAAGTIGLAVLRRLAPFDVHLHYDHRHLPSEVEKELNLTWHATR 241
 QY 243 EDMYPVCDVVTLCPLHPETEHMINDETLKFRRGAYIVNTARGKLCDRDAVARALESGR 302
 DB 242 EDMYPVCDVVTLCPLHPETEHMINDETLKFRRGAYIVNTARGKLCDRDAVARALESGR 301
 QY 303 LAGYAGDVWFPQAPADHPKRTMPYNGMTPHISGTLTAQARIAAGTRELCEFFGGRPI 362
 DB 302 LAGYAGDVWFPQAPADHPKRTMPYNGMTPHISGTLTAQARIAAGTRELCEFFGGRPI 361
 QY 363 RDEYLIYOGGALAGTGAHSHYSKGNATGSGSEAAKFKKAV 401
 DB 362 RDEYLIYOGGALAGTGAHSHYSKGNATGSGSEAAKFKKAV 400
 RESULT 2
 FDH_PICAN
 ID FDH_PICAN STANDARD; PRT; 361 AA.
 AC P33677;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, last sequence update)
 DT 28-FEB-2003 (Rel. 41, last annotation update)
 DE Formate dehydrogenase (EC 1.2.1.2) (NAD-dependent formate
 dehydrogenase) (FDH).
 GN PMDH.
 OS Pichia angusta (Yeast) (Hansenula polymorpha).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4905;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hollenberg C.P., Janowicz Z.;
 RT "DNA-molecules coding for PMDH control regions and structured gene for
 a protein having FMDH-activity and their uses.";
 RL Patent number EP0299108, 18-JAN-1989.
 CC -|- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
 CC -|- SUBUNIT: Homodimer.
 CC -|- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; A06214; CA00531.1; -
 DR HSSP; P33160; 2NAD.
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam; PF00389; 2-Hacid_DH.1.
 DR Pfam; PF02826; 2-Hacid_DH.C.1.
 DR PROSITE; PS00065; D_2_HYDROXYACID_DH.1; 1.
 DR PROSITE; PS00670; D_2_HYDROXYACID_DH.2; 1.
 DR PROSITE; PS00671; D_2_HYDROXYACID_DH.3; 1.
 KW Oxidoreductase; NAD.
 FT INT_MET 0
 FT NP_BIND 164 199 NAD (ADP PART) (BY SIMILARITY).
 FT ACT_SITE 257 257 SUBSTRATE BINDING (BY SIMILARITY).
 SQ SEQUENCE 361 AA; 39779 MW; 0FC001366F9E479B CRC64;
 Query Match 4.0%; Score 16; DB 1; Length 361;
 Best Local Similarity 100.0%; Pred. No. 2.8e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 308 GDVWFPPAPKDPHWR 323
 DB 280 GDVWFPPAPKDPHWR 295
 RESULT 3
 FDH_HORVU STANDARD; PRT; 377 AA.
 AC 09ZRI8;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Formate dehydrogenase, mitochondrial precursor (EC 1.2.1.2) (NAD-
 DE dependent formate dehydrogenase) (FDH).
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticaceae; Hordeum.
 OC NCBI_TaxID=4513;
 RN
 RP SEQUENCE FROM N.A.
 RP STRAIN=cv. Ehimedaka No.1; TISSUE=Root;
 RX MEDLINE=98150277; PubMed=9489019;
 RA Suzuki K., Itai R., Suzuki K., Nakanishi H., Nishizawa N.K.,
 RA Yoshimura E., Mori S.;
 RT "Formate dehydrogenase, an enzyme of anaerobic metabolism, is induced
 RT by iron deficiency in barley roots.";
 RL Plant Physiol. 116:725-732(1998).
 CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) -> CO(2) + NADH.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 CC DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D88272; BAA36181.1; -
 DR HSSP; P33160; 2NAD.
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam; PF00389; 2-Hacid_DH.1.
 DR Pfam; PF02826; 2-Hacid_DH.C.1.
 DR PROSITE; PS00065; D_2_HYDROXYACID_DH.1; FALSE_NEG.
 DR PROSITE; PS00670; D_2_HYDROXYACID_DH.2; 1.
 DR PROSITE; PS00671; D_2_HYDROXYACID_DH.3; 1.

KW Oxidoreductase; NAD; Mitochondrion; Transit peptide.
 FT TRANSIT 1 21. MITOCHONDRION (BY SIMILARITY).
 FT CHAIN 22 377 FORMATE DEHYDROGENASE.
 FT NP_BIND 191 225 NAD (ADP PART) (BY SIMILARITY).
 FT ACT_SITE 283 283 SUBSTRATE BINDING (BY SIMILARITY).
 SQ SEQUENCE 377 AA; 41545 MW; E28C1FE24E9225C7 CRC64;
 Query Match 4.0%; Score 16; DB 1; Length 377;
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 308 GDVWFPPAPKDPHWR 323
 DB 306 GDVWFPPAPKDPHWR 321
 RESULT 4
 FDH_SOLTU STANDARD; PRT; 381 AA.
 AC 007511; 09ZRI8;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Formate dehydrogenase, mitochondrial precursor (EC 1.2.1.2) (NAD-
 DE dependent formate dehydrogenase) (FDH).
 GN FDH1.
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
 OC NCBI_TaxID=4113;
 RN
 RP SEQUENCE FROM N.A.
 RP STRAIN=cv. BF15;
 RC Houton-Cabassa C., Ambard-Bretteville F., Remy R.,
 RA Colas des Francs-Small C.;
 RT "Evidence for multiple copies of formate dehydrogenase genes in
 RT plants: isolation of three potato fdh genes fdh1, fdh2 and fdh3.";
 RL (in) Plant Gene Register FGR98-102.
 RN
 RP SEQUENCE OF 3-381 FROM N.A., AND PARTIAL SEQUENCE.
 RP STRAIN=cv. BF15; TISSUE=Tuber;
 RX MEDLINE=94105343; PubMed=8278546;
 RA Colas des Francs-Small C., Ambard-Bretteville F., Small I.D.,
 RA Remy R.;
 RT "Identification of a major soluble protein in mitochondria from
 RT nonphotosynthetic tissues as NAD-dependent formate dehydrogenase.";
 RL Plant Physiol. 102:1171-1177(1993).
 RN
 RP REVISIONS TO N-TERMINUS.
 RC STRAIN=cv. BF15; TISSUE=Tuber;
 RA Colas des Francs-Small C.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE OF 26-54.
 RC STRAIN=cv. BF15; TISSUE=Tuber;
 RA Colas des Francs-Small C., Ambard-Bretteville F., Darpas A.,
 RA Sallantin M., Huet J.-C., Pernollet J.-C., Remy R.;
 RT "Variation of the polypeptide composition of mitochondria isolated
 RT from different potato tissues.";
 RL Plant Physiol. 98:273-278(1992).
 RN
 RP CHARACTERIZATION:
 RX PubMed=9490763;
 RA Houton-Cabassa C., Ambard-Bretteville F., Moreau F.,
 RA Davy de Virville J., Remy R., Colas des Francs-Small C.;
 RT "Stress induction of mitochondrial formate dehydrogenase in potato
 RT leaves.";
 RL Plant Physiol. 116:627-635(1998).
 CC -1- FUNCTION: Involved in formate-dependent oxygen uptake coupled to
 CC ATP synthesis.
 CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) -> CO(2) + NADH.
 CC -1- SUBUNIT: Homodimer.

```

CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- TISSUE SPECIFICITY: Found at high levels in developing tubers, at
CC intermediate level in stems, veins, stolons, and stamens,
CC and at low level in leaves and roots.
CC -1- INDUCTION: Induced very rapidly by wounding, and slower by
CC darkness, chilling, drought, hypoxia, and treatments with formate,
CC abscisic acid, serine, sarcosine, pyruvate, acetate, ethanol or
CC methanol.
CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.
CC -1- CAUTION: There are two other putative pseudogenes, fdh2 and fdh3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z59991; CAB17080.1; -
DR EMBL: Z59992; -; NOT_ANNOTATED_CDS.
DR EMBL: Z21493; CAA79702.2; -
DR PIR: J02272; J02272.
DR HSSP: P33160; 2NAD.
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR Pfam: PF00389; 2-Hacid_DH; 1.
DR Pfam: PF02826; 2-Hacid_DH.C; 1.
DR PROSITE: PS00065; D_2-HYDROXYACID_DH_1; 1.
DR PROSITE: PS00670; D_2-HYDROXYACID_DH_2; 1.
DR PROSITE: PS00671; D_2-HYDROXYACID_DH_3; 1.
DR Oxidoreductase; NAD; Mitochondrion; Transit peptide.
FT TRANSIT 1 25
FT CHAIN 1 25
FT NP_BIND 26 381
FT ACT_SITE 195 229
FT ACT_SITE 287 287
SQ SEQUENCE 381 AA; 42038 MW; 920D351AC5A3A00E CRC64;

Query Match
Best local Similarity 100.0%; Pred. No. 0.00037;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 LTAGISDHVDL 130
DB 121 LTAGISDHVDL 132

```

```

RX PubMed:12655011;
RA Mannhaupt G., Montrone C., Haase D., Mewes H.-W., Aln V.,
RA Hohnselt J.D., Fartmann B., Nykatura G., Kempken F., Maier J.,
RA Schulte U.;
RT "What's in the genome of a filamentous fungus? Analysis of the
RT Neurospora genome sequence.";
RL Nucleic Acids Res. 31:1944-1954(2003).
CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- DEVELOPMENTAL STAGE: DEVELOPMENTALLY REGULATED. EXPRESSED ONLY
CC DURING CONIDATION AND EARLY GERMINATION.
CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: I13964; AAA9900.1; -
DR EMBL: A451018; CAC18252.1; -
DR PIR: A47117; A47117.
DR HSSP: P33160; 2NAC.
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR Pfam: PF00389; 2-Hacid_DH; 1.
DR Pfam: PF02826; 2-Hacid_DH.C; 1.
DR PROSITE: PS00065; D_2-HYDROXYACID_DH_1; 1.
DR PROSITE: PS00670; D_2-HYDROXYACID_DH_2; 1.
DR PROSITE: PS00671; D_2-HYDROXYACID_DH_3; 1.
DR Oxidoreductase; NAD.
FT NP_BIND 166 201
FT ACT_SITE 259 259
FT DOMAIN 364 374
SQ SEQUENCE 375 AA; 40956 MW; 3073CB95F204968 CRC64;

Query Match
Best local Similarity 100.0%; Pred. No. 0.0038;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 TAGIGSDHVDL 130
DB 91 TAGIGSDHVDL 101

```

```

RESULT 5
FDH_NEUCR STANDARD; PRT; 375 AA.
ID FDH_NEUCR STANDARD; PRT; 375 AA.
AC 007103;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-FEB-2003 (Rel. 42, Last annotation update)
DE Formate dehydrogenase (EC 1.2.1.2) (NAD-dependent formate
DE dehydrogenase) (FDH).
GN FDH OR 99H12.160.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;
RX MEDLINE=93385982; PubMed=8509325;
RA Chow C.M., Rajbandary U.L.;
RT "Developmental regulation of the gene for formate dehydrogenase in
RT Neurospora crassa.";
RL J. Bacteriol. 175:3703-3709(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;

```

```

RESULT 6
FDH_ORYSA STANDARD; PRT; 376 AA.
ID FDH_ORYSA STANDARD; PRT; 376 AA.
AC 095XP2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Formate dehydrogenase, mitochondrial precursor (EC 1.2.1.2) (NAD-
DE dependent formate dehydrogenase) (FDH).
GN Oryza sativa (Rice).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriaraloideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Shiraiishi T., Fukusaki E., Kobayashi A.;
RT "NAD-dependent formate dehydrogenase.";
RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
CC -1- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB019533; BAA77337.1; -
 DR HSSP: P33160; 2NAD.
 DR Gramene; O9SKXP2; -
 DR InterPro: IPR006139; 2-Hacld_DH.
 DR InterPro: IPR006140; 2-Hacld_DH.C.
 DR Pfam: PF00389; 2-Hacld_DH; 1.
 DR Pfam: PF02826; 2-Hacld_DH.C; 1.
 DR PROSITE: PS00665; D_2-HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D_2-HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2-HYDROXYACID_DH_3; 1.
 DR Oxidoreductase; NAD; Mitochondrion; Transit peptide.
 DR TRANSIT 1 20 MITOCHONDRION (BY SIMILARITY).
 FT CHAIN 21 376 FORMATE DEHYDROGENASE.
 FT NP_BIND 190 224 NAD (ADP PART) (BY SIMILARITY).
 FT ACT_SITE 282 282 SUBSTRATE BINDING (BY SIMILARITY).
 SO SEQUENCE 376 AA; 41243 MW; 9D3964B64C2F4040 CRC64;
 Query Match 2.7%; Score 11; DB 1; Length 376;
 Best Local Similarity 100.0%; Pred. No. 0.0038;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 308 GDVWFPQAPK 318
 Db 305 GDVWFPQAPK 315
 RESULT 7
 FDH_EMENI STANDARD: PRT; 377 AA.
 AC Q03134;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable formate dehydrogenase (EC 1.2.1.2) (NAD-dependent formate
 DE dehydrogenase) (FDH).
 GN ACIA.
 OS *Emricella nidulans* (*Aspergillus nidulans*).
 CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 CC Eurotiales; Trichocomaceae; *Emricella*.
 CC NCBL_TaxID=162425;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93101140; PubMed-1465107;
 RA Saleeba J.A., Cobbett C.S., Hynes M.J.;
 RT "Characterization of the amda-regulated acia gene of *Aspergillus*
 RT *nidulans*."; Mol. Gen. Genet. 235:349-358(1992).
 CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- INDUCTION: ACETATE INDUCTION MEDIATED BY AMDA REGULATORY GENE.
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 CC DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z11612; CA77687.1; -
 DR PIR: S30088; S30088.
 DR HSSP: P33160; 2NAC.
 DR InterPro: IPR006139; 2-Hacld_DH.
 DR InterPro: IPR006140; 2-Hacld_DH.C.

DR Pfam: PF00389; 2-Hacld_DH; 1.
 DR Pfam: PF02826; 2-Hacld_DH.C; 1.
 DR PROSITE: PS00665; D_2-HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D_2-HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2-HYDROXYACID_DH_3; FALSE-NEG.
 DR Oxidoreductase; NAD.
 FT NP_BIND 161 191 NAD (ADP PART) (BY SIMILARITY).
 FT ACT_SITE 271 271 SUBSTRATE BINDING (BY SIMILARITY).
 SO SEQUENCE 377 AA; 41531 MW; 5219F5F8CC931CE CRC64;
 Query Match 2.7%; Score 11; DB 1; Length 377;
 Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 120 TAGIGSDHVDL 130
 Db 86 TAGIGSDHVDL 96
 RESULT 8
 FDH_ARATH STANDARD: PRT; 384 AA.
 AC Q957E4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Formate dehydrogenase, mitochondrial precursor (EC 1.2.1.2) (NAD-
 DE dependent formate dehydrogenase) (FDH).
 GN FDH1 OR FDH OR AT5G14780 OR T9L3_80.
 OS *Arabidopsis thaliana* (Mouse-ear cress).
 CC *Arabidopsis thaliana* (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eustosids II; Brassicales; Brassicaceae; *Arabidopsis*.
 CC NCBL_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nishikawa T., Fukusaki E., Kobayashi A.;
 RT "Formate dehydrogenase cDNA from *Arabidopsis thaliana*."; Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Li R., Bonham-Smith P.C., King J.;
 RT "Arabidopsis thaliana mRNA for NAD-dependent formate dehydrogenase
 RT 1."; Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Markwell J., Osterman J.C., Olson B.J., Skavdahl M., Ramberg H.,
 RA Germann M.C.;
 RT "Induction of leaf formate dehydrogenase by one-carbon metabolites."; Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE-21016721; PubMed-11130714;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsuno M., Matsuno A., Muraki A., Nakayama S.,
 RA Nakazaki N., Nario K., Okumura S., Shilpo S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Sato M., de la Bastide M.,
 RA Huang E., Spiegel L., Gao J., O'Shaughnessy A., Preston R.,
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
 RA Stoneking T., Pepin K., Spieth J., Sakhon M., Armstrong J., Becker M.,
 RA Belter E., Cordum H., Cordes M., Courtney W., Dante M.,
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
 RA Leonard S., Meyer R., Mulvaney E., Ozeraky P., Riley A., Stromwall C.,
 RA Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,
 RA Parrell L., Shah R., Rodriguez M., Hoon See L., Vill D., Baker J.,
 RA Kirchhoff K., Toth K., King L., Bahret A., Miller B., Marra M.,
 RA Martensen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Volckaert G., Wandelt R., Duesterhoelt A., Stiekema W., Pohl T.,
 RA Entian K.-D., Terryn N., Hartley N., Bent E., Johnson S.,
 RA Langham S.-A., McCullagh B., Robben J., Grynprenzer B., Zimmermann W.,

RA Ramsperger U., Wedler H., Balke K., Wedler E., Peters S.,
 RA van Staveren M., Dirke W., Moolman P., Klein Lankhorst R.,
 RA Weltzenger T., Botke G., Rose M., Haut J., Berner S.,
 RA Feldpausch M., Lamberth S., Villarroel R., Giesen J., Adies W.,
 RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.,
 RA "Sequence and analysis of chromosome 5 of the plant Arabidopsis
 thaliana."
 RT Nature 408:823-826(2000).
 RL [5]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, Columbia;
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "Riken Arabidopsis full length cDNA clones (RAFLs) sequenced by the
 RT SGP consortium (Salk/Stanford/PGSC).";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- CAPABILITY ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 CC DEHYDROGENASES FAMILY. STRONG, TO OTHER FGH.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB023897; BAB86683.1; -
 DR EMBL: AF208028; AAF19435.1; -
 DR EMBL: AF208029; AAF19436.1; -
 DR EMBL: AF217195; AAF67100.1; -
 DR EMBL: AL391149; AAC01877.1; -
 DR EMBL: AY054285; AAL06944.1; -
 DR EMBL: AY039609; AAK62664.1; -
 DR EMBL: AY081734; AAL87387.1; -
 DR F1R; T51423; T51423.
 DR HSP; P33160; 2NAD.
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH_C.
 DR Pfam: PF02836; 2-Hacid_DH_1.
 DR Pfam: PF02836; 2-Hacid_DH_C; 1.
 DR PROSITE: PS00065; D-2-HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D-2-HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D-2-HYDROXYACID_DH_3; FALSE_NEG.
 KM Oxidoreductase; NAD; Mitochondrion; Transit peptide.
 FT TRANSIT 1 27
 FT CHAIN 28 384
 FT NE_BIND 198 232
 FT ACT_SITE 290 290
 SQ SEQUENCE 384 AA; 42409 MW; A12BA423019D862B CRC64;
 Query Match 2.7%; Score 11; DB 1; Length 384;
 Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 313 POPAPKDPWR 323
 Db 318 POPAPKDPWR 328
 RESULT 9
 FRDA_DROME STANDARD; PRT; 190 AA.
 AC Q9W385;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Frataxin homolog, mitochondrial precursor.
 GN FH OR CG8971.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20509989; PubMed=11054533;
 RA Canizares J., Blanca J.M., Navarro J.A., Monros E., Palau F.,
 RA Molto M.D.;
 RT "dth is a Drosophila homologue of the Friedreich's ataxia disease
 RT gene".
 RL Gene 256:35-42(2000).
 RN
 RP
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V., Chapple M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berna B.P., Bhandal D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brocksen P., Brotler P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrelle S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hristin D., Houston K.A., Howard T.J., Wei M.-H., Idegam C.,
 RA Jhalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Snyder E.C., Spredell A.C., Stapleton M., Strong R., Sun E.,
 RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: PROBABLY INVOLVED IN IRON HOMEOSTASIS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE FRATAXIN FAMILY.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF208492; AAG35733.1; -
 DR EMBL: AE003446; AAF46449.1; -
 DR HSP; Q16595; 1DLX.
 DR FlyBase: FBgn0030092; fh.
 DR InterPro: IPR001794; Frataxin.
 DR InterPro: IPR002908; Frataxin_like.
 DR Pfam: PF01491; Frataxin_Cyay; 1.
 DR PRINTS; PR00904; FRATAXIN.

DR PRODOM: PD238818; Frataxin_1like; 1.
 DR PROSITE: PS01344; FRATAXIN_1; 1.
 DR PROSITE: PS0810; FRATAXIN_2; 1.
 KW Mitochondrion; Transit peptide.
 FT TRANSIT 1 190 MITOCHONDRION.
 FT CHAIN ? FRATAXIN HOMOLOG.
 SQ SEQUENCE 190 AA; 20921 MM; 8C48332C39F3036 CRC64;

Query Match 2.0%; Score 8; DB 1; Length 190;
 Best Local Similarity 100.0%; Pred. No. 2.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 195 VGTVAAGR 202
 DB 147 VGTVAAGR 154

RESULT 10
 Y106_MYCTU STANDARD; PRT; 398 AA.
 AC Q10899;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein RV0106.
 GN RV0106 OR MT0115 OR MTCY251.25.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37Rv;
 RA MEDLINE=96295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badoock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean R., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Ruster S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sultun J.E., Taylor K., Whitehead S., Barrett B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
 CC -I- SIMILARITY: SOME. TO P.DENTRIFICANS COBW.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: Z74410; CA98919.1; -
 DR EMBL: AE006923; AAK4337.1; -
 DR PIR: B70752; B70752.
 DR TIGR: MT0115; -
 DR Tuberculist; RV0106; -
 DR InterPro; IPR003495; COBW.
 DR Pfam; PF02492; COBW; 1.

KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 398 AA; 43699 MM; 2D75D9FBA167054A CRC64;

Query Match 2.0%; Score 8; DB 1; Length 398;
 Best Local Similarity 100.0%; Pred. No. 4.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 206 AVLRRLAP 213
 DB 181 AVLRRLAP 188

RESULT 11
 PURA_XANAC STANDARD; PRT; 430 AA.
 AC Q8PNB5;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Adenylosuccinate synthetase (EC 6.3.4.4) (IMP--aspartate ligase)
 DE (AdSS) (AMPase).
 GN PURA OR XAC1158.
 OS Xanthomonas axonopodis (pv. citri).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=92829;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=306 / ATCC 13902 / XV 101;
 RA MEDLINE=22022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furian L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Canavann F., Cardozo J., Chambergo F., Clapira L.P.,
 RA Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Forlignieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locall E.C., Machado M.A., Madelira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Medeiros J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitejima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities."
 RL Nature 417:459-463(2002).
 CC -I- FUNCTION: Plays an important role in the de novo pathway of purine
 CC nucleotide biosynthesis.
 CC -I- CATALYTIC ACTIVITY: GTP + IMP + L-aspartate = GDP + phosphate +
 CC adenylosuccinate.
 CC -I- PATHWAY: AMP biosynthesis; first committed step.
 CC -I- SIMILARITY: Belongs to the adenylosuccinate synthetase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AE011745; AAM36030.1; -
 DR HAMAP; MF_00011; -
 DR InterPro; IPR001114; Asucc_synthase.
 DR Pfam; PF00709; Adenylosucc_synth. 1.
 DR PRODOM; PD001188; Asucc_synthase; 1.
 DR TIGR; TIGR0184; purA; 1.
 DR PROSITE; PS01266; ADENYLOSUCCIN_SYN_1; 1.
 DR PROSITE; PS00513; ADENYLOSUCCIN_SYN_2; 1.
 DR Purine biosynthesis; Ligase; GTP-binding; Complete proteome.
 FT NP_BIND 13 19 GTP (POTENTIAL).

```

FT ACT_SITE 141 141 BY SIMILARITY.
FT ACT_SITE 148 148 BY SIMILARITY.
SQ SEQUENCE 430 AA; 46199 MW; 770E439BE1398F8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 1; Length 430;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 371 GGALACTG 378
Db 247 GGALACTG 254

RESULT 12
PURA_XANCP
ID PURA_XANCP STANDARD; PRT; 430 AA.
AC 08PBB6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Adenylosuccinate synthetase (EC 6.3.4.4) (IMP--aspartate ligase)
DE (AdSS) (AMPase).
GN PURA OR XC01053.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_Taxid=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 33913 / NCPPB 528;
RX MEDLINE-22022145; PubMed-12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Alves L.M.C., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Caccarette G., Camargo F., Cardozo J., Chambeiro F., Clapina L.P.,
RA Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Baria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsal S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).

-1- FUNCTION: Plays an important role in the de novo pathway of purine
nucleotide biosynthesis.
-1- CATALYTIC ACTIVITY: GTP + IMP + L-aspartate -> GDP + phosphate +
adenylosuccinate.
-1- PATHWAY: AMP biosynthesis; first committed step.
-1- SIMILARITY: Belongs to the adenylosuccinate synthetase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
CC EMBL; AEO12202; AAM40352.1;
CC
CC HAMAP; MF_00011; 1.
CC InterPro; IPR001114; Asucc_synthase.
CC Pfam; PF00709; Adenylosucc_synth; 1.
CC ProDom; PD001188; Asucc_synthase; 1.
CC TRFAMS; TIGR00184; PURA; 1.
CC PROSITE; PS01266; ADENYLOSUCCLIN_SYN_1; 1.
CC PROSITE; PS00513; ADENYLOSUCCLIN_SYN_2; 1.
CC Purine biosynthesis; Ligase; GTP-binding; Complete proteome.
KW

```

```

FT NP_BIND 13 19 GTP (POTENTIAL).
FT ACT_SITE 141 141 BY SIMILARITY.
FT ACT_SITE 148 148 BY SIMILARITY.
SQ SEQUENCE 430 AA; 46163 MW; D36D57EAC2081B2D CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 1; Length 430;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 371 GGALACTG 378
Db 247 GGALACTG 254

RESULT 13
PFDB_SUISO
ID PFDB_SUISO STANDARD; PRT; 126 AA.
AC 09UXB8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Prefoldin beta subunit (GlmC beta subunit).
DE PFDB OR SS00730 OR C20_021.
GN Sulfolobus solfataricus.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_Taxid=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35092 / DSM 1617 / P2;
RX MEDLINE-20165948; PubMed-10701121;
RA Charlebois R.L., Singh R.K., Chan-Weher C.C.-Y., Allard G., Chow C.,
RA Confalonieri F., Curtis B., Duguet M., Eruso G., Faguy D.,
RA Gaasterland T., Garrett R.A., Gordon P., Jeffries A.C., Kozera C.,
RA Kushwaha N., Lafleur E., Medina N., Peng X., Penny S.L., She Q.,
RA St Jean A., van der Oost J., Young F., Zivanovic T., Doolittle W.F.,
RA Ragan M.A., Senses C.W.;
RT "gene content and organization of a 281-kbp contig from the genome of
RT the extremely thermophilic archaeon, Sulfolobus solfataricus P2."
RL Genome 43:116-136(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35092 / DSM 1617 / P2;
RX MEDLINE-21332296; PubMed-11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweyer M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moers A., Eruso G., Fletcher C., Gordon P.M.K.,
RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
RT Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
RL
-1- FUNCTION: Molecular chaperone capable of stabilizing a range of
proteins. Seems to fulfil an ATP-independent, HSP70-like function
in archaeal de novo protein folding (by similarity).
-1- SUBUNIT: Heterohexamer of two alpha and four beta subunits (By
similarity).
-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
-1- SIMILARITY: BELONGS TO THE PREFOLDIN BETA SUBUNIT FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
CC EMBL; Y18930; CAB57573.1;
CC
CC EMBL; AEO06598; AAK41027.1;
CC
CC PIR; D90221; D90221.
CC HSP; O26774; 1FXK.

```

```

DR HAMAP: MF_00307; -: 1.
DR InterPro: IPR002777; PrefoldinKE2.
DR Pfam: PF01920; KE2; 1.
KW Chaperone; Complete proteome.
SO SEQUENCE 126 AA; 14543 MW; 5895DB741C8DE117 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 229 SVEKELN 235
DB 71 SVEKELN 77

RESULT 14
ID Y605_METUA STANDARD; PRT; 134 AA.
AC Q58022;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein M0605.
GN M0605.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
CX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=9633799; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Claydon R.A., Gocayne J.D.,
RA Ketravage A.N., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii."
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: STRONG, TO M.JANNASCHII M0142.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U67508; AAB98599.1; -.
CC DR TIGR: MJ0605; -.
CC DR InterPro: IPR002806; DUF103.
CC DR Pfam: PF01953; DUF103; 1.
CC DR ProDom: PD015886; DUF103; 1.
CC KW Hypothetical protein; Complete proteome.
CC SEQUENCE 134 AA; 15596 MW; 4C1D5D528BD3452C CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 1; Length 134;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 291 RDAVARA 297
DB 37 RDAVARA 43

RESULT 15
UCN3_HUMAN

```

```

ID UCN3_HUMAN STANDARD; PRT; 161 AA.
AC Q969E3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urocortin III precursor (Ucn III) (Stresscopin).
GN UCN3 OR SPC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SYNTHESIS OF UCN III.
RX MEDLINE=21310014; PubMed=11416224;
RA Lewis K., Li C., Perrin M.H., Blount A., Kunitake K., Donaldson C.,
RA Vaughan J., Reyes T.M., Guylas J., Fischer W., Bilezikjian L.,
RA Rivier J., Sawchenko P.E., Vale W.W.;
RT "Identification of urocortin III, an additional member of the
RT corticotropin-releasing factor (CRF) family with high affinity for the
RT CRF receptor."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7570-7575(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21227098; PubMed=11329063;
RA Hsu S.Y., Hsueh A.J.W.;
RT "Human stresscopin and stresscopin-related peptide are selective
RT ligands for the type 2 corticotropin-releasing hormone receptor."
RL Nat. Med. 7:605-611(2001).
CC -1- FUNCTION: Suppresses food intake, delays gastric emptying and
CC decreases heat-induced edema. Might represent an endogenous ligand
CC for maintaining homeostasis after stress.
CC -1- SUBUNIT: Binds with high affinity to CRF receptors 2-alpha and 2-
CC beta.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SAVVAGINE/CORTICOTROPIN-RELEASING
CC FACTOR/UCROTENSIN I FAMILY OF PEPTIDES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF361943; AAK67317.1; -.
CC DR EMBL: AY026949; AAK11729.1; -.
CC DR MIM: 605901; -.
CC KW Hormone; Amidation; Signal.
CC FT SIGNAL. 1 21
CC FT PROPEP 22 118
CC FT CHAIN 120 157
CC FT MOD_RES 157 157
CC FT AMIDATION (G-158 PROVIDE AMIDE GROUP).
CC SEQUENCE 161 AA; 17862 MW; F92B2C8C3087857D CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 109 IAKAKNL 115
DB 137 IAKAKNL 143

Search completed: July 24, 2003, 20:59:36
Job time : 26 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:56:19 ; Search time 99 Seconds
(without alignments)
1045.243 Million cell updates/sec

Title: US-09-996-008b-2
Perfect score: 401
Sequence: 1 MAKVLCVLYDDPVDGYPKTY.....YSKGNATGSGSEAKFKKAV 401

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	339	84.5	400	2	Q9R5V0
2	339	84.5	401	2	Q9R5V0 mycobacteri
3	31	7.7	402	2	Q9R5V1 mycobacteri
4	24	6.0	401	16	Q9R5V1 moraxella s
5	23	5.7	400	2	Q9R5V3 rhizobium m
6	20	5.0	399	2	Q9R5V3 paracoccus
7	20	4.0	248	10	Q9R5V1 hyphomicrob
8	16	4.0	364	3	Q9R5V3 pinus pinas
9	16	4.0	364	3	Q9R5V3 candida boi
10	16	4.0	364	3	Q9R5V3 candida met
11	13	3.0	398	2	Q9R5V3 candida boi
12	13	3.0	398	2	Q9R5V3 candida boi
13	12	3.0	376	3	Q9R5V3 candida boi
14	11	2.7	152	16	Q9R5V3 candida boi
15	11	2.7	199	10	Q9R5V3 candida boi
16	11	2.7	417	3	Q9R5V3 candida boi

17	8	2.0	89	17	Q9R5V3	097YX3
18	8	2.0	147	2	Q9R5V3	096G07
19	8	2.0	164	2	Q9R5V3	056431
20	8	2.0	320	16	Q9R5V3	08D4V2
21	8	2.0	334	16	Q9R5V3	08UJ61
22	8	2.0	334	16	Q9R5V3	092T15
23	8	2.0	345	12	Q9R5V3	08GHS3
24	8	2.0	374	16	Q9R5V3	099X44
25	8	2.0	374	16	Q9R5V3	08N1M1
26	8	2.0	382	2	Q9R5V3	09K4W5
27	8	2.0	501	17	Q9R5V3	08ZTM7
28	8	2.0	616	16	Q9R5V3	08DSJ9
29	8	2.0	617	16	Q9R5V3	097SR1
30	8	2.0	617	16	Q9R5V3	08E362
31	8	2.0	617	16	Q9R5V3	08DXD8
32	8	2.0	617	16	Q9R5V3	08DRB0
33	8	2.0	618	16	Q9R5V3	099X44
34	8	2.0	618	16	Q9R5V3	08N2B4
35	8	2.0	618	16	Q9R5V3	08K557
36	8	2.0	758	16	Q9R5V3	09JZM9
37	8	2.0	764	16	Q9R5V3	09JUS2
38	8	2.0	849	15	Q9R5V3	08UT64
39	8	2.0	991	17	Q9R5V3	08ZVU0
40	8	2.0	998	17	Q9R5V3	08ZXL7
41	8	2.0	1105	16	Q9R5V3	09A430
42	8	2.0	1184	16	Q9R5V3	09F2P1
43	8	1.7	40	11	Q9R5V3	09QV65
44	7	1.7	77	2	Q9R5V3	0932W3
45	7	1.7	77	2	Q9R5V3	08RLN0

ALIGNMENTS

RESULT 1

ID	Q9R5V0	PRELIMINARY:	PRT:	400 AA.
AC	Q9R5V0:	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-OCT-2002 (TREMBLrel. 22, Last annotation update)			
DE	NAD(+)-dependent formate dehydrogenase (EC 1.2.1.2) (Fragment).			
OS	Mycobacterium vaccae.			
OC	Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;			
OC	Coriobacteriales; Mycobacteriaceae; Mycobacterium.			
OX	NCBI_TaxID=1810;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96158475; PubMed=8597552;			
RA	Galkin A., Kulakova L., Tishkov V., Esaki N., Soda K.;			
RT	"Cloning of formate dehydrogenase gene from a methanol-utilizing			
RT	bacterium Mycobacterium vaccae N10."			
RL	Appl. Microbiol. Biotechnol. 44:479-483(1995).			
DR	HSSP: P33160; 2NAD.			
DR	InterPro: IPR006139; 2-Hacid_DH.			
DR	InterPro: IPR006140; 2-Hacid_DH_C.			
DR	Pfam: PF00389; 2-Hacid_DH; 1.			
DR	Pfam: PF02826; 2-Hacid_DH_C; 1.			
DR	PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.			
DR	PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.			
DR	PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.			
SO	SEQUENCE 400 AA; 44018 MW; 4422888CA22BD9B CRC64;			

Query Match

Best Local Similarity 84.5%; Score 339; DB 2; Length 400;
Matches 339; Conservative 100.0%; Pred. No. 0;
Mismatches 0; Indels 0; Gaps 0;

QY	63	YLESNGHLYVTSKDGSDSVFERELNADAVYISQPEMPAYLTTPRIAKAKLALTAG	122
DB	62	YLESNGHLYVTSKDGSDSVFERELNADAVYISQPEMPAYLTTPRIAKAKLALTAG	121
QY	123	IGSDHVDLSAIDRNVTAEVYCNISVAEHVMMILSVNLYPSHEMARKGMNAD	182

```

Db      122  IGSDDHVDLQSAIDRNTVAETVYCNCSISVAEHVVMILSLVRNLTSPSEHARKGNMIAD 181
Qy      183  CVSHAYDLEAMHGTVAAGRIGLAVLRRLAPFDVHLHYTRDHRRLPESVEKELNLTWATR 242
Db      182  CVSHAYDLEAMHGTVAAGRIGLAVLRRLAPFDVHLHYTRDHRRLPESVEKELNLTWATR 241
Qy      243  EDMYPCDVVTLNCPRLPETHMINDETLTKFRGAVIYNTARGKLCDDRAVVAALLESGR 302
Db      242  EDMYPCDVVTLNCPRLPETHMINDETLTKFRGAVIYNTARGKLCDDRAVVAALLESGR 301
Qy      303  LAGYAGVMPFPQAPKDPKHPRTMPYNGMTPHISGTTTLTAQARYAAGTRELIECFEGRPI 362
Db      302  LAGYAGVMPFPQAPKDPKHPRTMPYNGMTPHISGTTTLTAQARYAAGTRELIECFEGRPI 361
Qy      363  RDEYLIYOGGALAGTGAHYSKGNATGSEBAKFKKAV 401
Db      362  RDEYLIYOGGALAGTGAHYSKGNATGSEBAKFKKAV 400

```

RESULT 2

```

Q93GV1  PRELIMINARY; PRT; 401 AA.
AC      093GV1:
DT      01-DEC-2001 (TREMblrel. 19, Created)
DT      01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT      01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE      Formate dehydrogenase (EC 1.2.1.2).
GN      FDH.
OS      Mycobacterium vaccae.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxID=1810;
RN      [1]
RP      Galkin A., Kulakova L., Tishkov V., Esaki N., Soda K.;
RT      Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=96158475; PubMed=8597552;
RA      Galkin A., Kulakova L., Tishkov V., Esaki N., Soda K.;
RT      "Cloning of formate dehydrogenase gene from a methanol-utilizing
RT      bacterium Mycobacterium vaccae N10."
RL      Appl. Microbiol. Biotechnol. 44:479-483(1995).
DR      EMBL; AB072394; BAB69476.1; -.
DR      InterPro; IPR006139; 2-Hacid_DH.
DR      InterPro; IPR006140; 2-Hacid_DH_C.
DR      Pfam; PF00389; 2-Hacid_DH_C.1.
DR      Pfam; PF02826; 2-Hacid_DH_C.1.
DR      PROSITE; PS00065; D_2_HYDROXYACID_DH_1; 1.
DR      PROSITE; PS00670; D_2_HYDROXYACID_DH_2; 1.
DR      PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.
KW      Oxidoreductase.
SQ      SEQUENCE 401 AA; 44149 MW; 8C4821B4790AAD26 CRC64;

```

```

Query Match      84.5%; Score 339; DB 2; Length 401;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      63  YLENSGHTLVVTSKDKDPSVFERELVDADVISOFPWPAYLTPERTAKAKNLTATAG 122
Db      63  YLENSGHTLVVTSKDKDPSVFERELVDADVISOFPWPAYLTPERTAKAKNLTATAG 122
Qy      123  IGSDDHVDLQSAIDRNTVAETVYCNCSISVAEHVVMILSLVRNLTSPSEHARKGNMIAD 182
Db      123  IGSDDHVDLQSAIDRNTVAETVYCNCSISVAEHVVMILSLVRNLTSPSEHARKGNMIAD 182
Qy      183  CVSHAYDLEAMHGTVAAGRIGLAVLRRLAPFDVHLHYTRDHRRLPESVEKELNLTWATR 242
Db      183  CVSHAYDLEAMHGTVAAGRIGLAVLRRLAPFDVHLHYTRDHRRLPESVEKELNLTWATR 242
Qy      243  EDMYPCDVVTLNCPRLPETHMINDETLTKFRGAVIYNTARGKLCDDRAVVAALLESGR 302
Db      243  EDMYPCDVVTLNCPRLPETHMINDETLTKFRGAVIYNTARGKLCDDRAVVAALLESGR 302

```

```

Qy      303  LAGYAGVMPFPQAPKDPKHPRTMPYNGMTPHISGTTTLTAQARYAAGTRELIECFEGRPI 362
Db      303  LAGYAGVMPFPQAPKDPKHPRTMPYNGMTPHISGTTTLTAQARYAAGTRELIECFEGRPI 362
Qy      363  RDEYLIYOGGALAGTGAHYSKGNATGSEBAKFKKAV 401
Db      363  RDEYLIYOGGALAGTGAHYSKGNATGSEBAKFKKAV 401

```

RESULT 3

```

Q08375  PRELIMINARY; PRT; 402 AA.
AC      008375:
DT      01-JUL-1997 (TREMblrel. 04, Created)
DT      01-JUL-1997 (TREMblrel. 04, Last sequence update)
DT      01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE      NAD-dependent formate dehydrogenase (EC 1.2.1.2).
GN      FDH.
OS      Moraxella sp.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC      Moraxellaceae; Moraxella.
OX      NCBI_TaxID=479;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Galkin A.G.;
RT      Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR      EMBL; Y13245; CAAT3696.1; -.
DR      HSSP; P33160; ZNAD.
DR      InterPro; IPR006139; 2-Hacid_DH.
DR      InterPro; IPR006140; 2-Hacid_DH_C.
DR      Pfam; PF00389; 2-Hacid_DH_C.1.
DR      Pfam; PF02826; 2-Hacid_DH_C.1.
DR      PROSITE; PS00065; D_2_HYDROXYACID_DH_1; 1.
DR      PROSITE; PS00670; D_2_HYDROXYACID_DH_2; 1.
DR      PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.
KW      NAD; Oxidoreductase.
SQ      SEQUENCE 402 AA; 44143 MW; C511CE1CE96BBB CRC64;

```

```

Query Match      7.7%; Score 31; DB 2; Length 402;
Best Local Similarity 100.0%; Pred. No. 5; 5e-23;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      249  CDVYTLNCPRLPETHMINDETLTKFRGAV 279
Db      249  CDVYTLNCPRLPETHMINDETLTKFRGAV 279

```

RESULT 4

```

Q930E7  PRELIMINARY; PRT; 401 AA.
AC      0930E7:
DT      01-DEC-2001 (TREMblrel. 19, Created)
DT      01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT      01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE      Probable NAD-dependent formate dehydrogenase.
GN      RA0251 OR SMA0478.
OS      Rhizobium meliloti (Sinorhizobium meliloti).
OC      Plasmid pSymA (megaplasmid 1).
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Rhizobiaceae; Sinorhizobium.
OX      NCBI_TaxID=382;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=1021;
RX      MEDLINE=21396509; PubMed=11481432;
RA      Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA      Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA      Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA      Kallman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA      Yeh K.-C., Davis R.W., Federici N.A., Long S.R.;
RT      "Nucleotide sequence and predicted functions of the entire

```

RT Shnorhizobium meliloti pSyma megaplasmid.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
 DR EMBL: AE007218; AAK64909.1; -
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam: PF00389; 2-Hacid_DH; 1.
 DR Pfam: PF02826; 2-Hacid_DH.C; 1.
 DR PROSITE: PS00065; D_2_HYDROXYACID_DH.1; 1.
 DR PROSITE: PS00670; D_2_HYDROXYACID_DH.2; 1.
 DR PROSITE: PS00671; D_2_HYDROXYACID_DH.3; 1.
 KW Plasmid; Complete proteome.
 SQ SEQUENCE 401 AA; 43533 MW; C2693A417456204B CRC64;

Query Match 6.0%; Score 24; DB 16; Length 401;
 Best Local Similarity 100.0%; Pred. No. 9.4e-16;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 GHTLVVTSKDKGPDSPFERELVDA 91
 ||||||||||||||||||
 DB 70 GHTLVVTSKDKGPDSPFERELVDA 93

RESULT 5
 OY3GW3 PRELIMINARY; PRT; 400 AA.
 AC OY3GW3;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE NAD-dependent formate dehydrogenase.
 GN FDH.
 OS Paracoccus sp. 12-A.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
 CC Rhodobacteraceae; Paracoccus.
 OX NCBI_TaxID=171192;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-12-A;
 RA Shinoda T., Satoh T., Mineki S., Iida M., Taguchi H.;
 RT "Cloning, nucleotide sequence and expression in Escherichia coli of
 the gene for formate dehydrogenase of Paracoccus sp. 12-A, a formate-
 assimilating bacterium.";
 RL Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB071373; BAB64941.1; -
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam: PF00389; 2-Hacid_DH; 1.
 DR Pfam: PF02826; 2-Hacid_DH.C; 1.
 DR PROSITE: PS00065; D_2_HYDROXYACID_DH.1; 1.
 DR PROSITE: PS00670; D_2_HYDROXYACID_DH.2; 1.
 DR PROSITE: PS00671; D_2_HYDROXYACID_DH.3; 1.
 SQ SEQUENCE 400 AA; 43757 MW; 13195AED9AC299D8 CRC64;

Query Match 5.7%; Score 23; DB 2; Length 400;
 Best Local Similarity 100.0%; Pred. No. 1e-14;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 378 GAHSYSKGNATGSEBAKFKKA 400
 ||||||||||||||||||
 DB 378 GAHSYSKGNATGSEBAKFKKA 400

RESULT 6
 OY3UW1 PRELIMINARY; PRT; 399 AA.
 AC OY3UW1;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE NAD+-dependent formate dehydrogenase.
 GN FDH.
 OS Hyphomicrobium sp. JC17.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Hyphomicrobiaceae; Hyphomicrobium.
 OX NCBI_TaxID=142666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JC17;
 RA Tanaka Y., Yoshida T., Watanabe K., Mitsuana T.;
 RT "Characterization, gene cloning and expression of NAD+-dependent
 formate dehydrogenase from a methylotrophic bacterium Hyphomicrobium
 sp. JC17.";
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB051073; BAB55449.1; -
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam: PF00389; 2-Hacid_DH; 1.
 DR Pfam: PF02826; 2-Hacid_DH.C; 1.
 DR PROSITE: PS00671; D_2_HYDROXYACID_DH.3; 1.
 SQ SEQUENCE 399 AA; 44044 MW; 853BAD2C635864DC CRC64;

Query Match 5.0%; Score 20; DB 2; Length 399;
 Best Local Similarity 100.0%; Pred. No. 1.3e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 249 CDVYTLNCPHPETEEMIND 268
 ||||||||||||||||||
 DB 249 CDVYTLNCPHPETEEMIND 268

RESULT 7
 OYVX85 PRELIMINARY; PRT; 248 AA.
 AC OYVX85;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Putative NAD-dependent formate dehydrogenase (EC 1.2.1.2)
 (Fragment).
 OS Pinus pinaster (Maritime pine).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
 OX NCBI_TaxID=71647;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Root;
 RA Dubos C., Plomion C.;
 RT "Identification of water-deficit responsive genes in maritime pine
 (Pinus pinaster Ait.) roots using non radioactive cDNA-AFLP.";
 RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AJ300721; CAC83306.1; -
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam: PF02826; 2-Hacid_DH.C; 1.
 DR PROSITE: PS00065; D_2_HYDROXYACID_DH.1; 1.
 DR PROSITE: PS00670; D_2_HYDROXYACID_DH.2; 1.
 DR PROSITE: PS00671; D_2_HYDROXYACID_DH.3; 1.
 KW Oxidoreductase; NAD.
 RN NON_TER
 FT NON_TER
 SQ SEQUENCE 248 AA; 27370 MW; 45E61CFAPD55A0C8 CRC64;

Query Match 4.0%; Score 16; DB 10; Length 248;
 Best Local Similarity 100.0%; Pred. No. 1.1e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 308 GDVWFPDPAPKDPWR 323
 ||||||||||||||||||
 DB 174 GDVWFPDPAPKDPWR 189

RESULT 8
 OY3437 PRELIMINARY; PRT; 364 AA.
 AC OY3437;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)

```

DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE NAD-dependent formate dehydrogenase (EC 1.2.1.2).
GN FPH1 OR FDH3.
OS Candida boidinii (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
ON NCBI_TaxID=5477;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=52;
RX MEDLINE=97369815; PubMed=9226256;
RA Sakai Y., Murdanoto A.P., Konishi T., Iwamatsu A., Kato N.;
RT "Regulation of the formate dehydrogenase gene, FPH1, in the
RT methylophilic yeast Candida boidinii and growth characteristics of an
RT FPH1-disrupted strain on methanol, methylamine, and choline.";
RL J. Bacteriol. 179:4480-4485(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC32195;
RA Siusarczyk H.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC32195;
RA Siusarczyk H.;
RL Theiss (1997), Heinrich-Heine-University Duesseldorf.
DR EMBL: AF004096; AAC49766.1; -.
DR EMBL: AJ245934; CAB54834.1; -.
DR HSSP: P33160; 2NAC.
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR Pfam: PF00389; 2-Hacid_DH.1.
DR Pfam: PF02826; 2-Hacid_DH.C.1.
DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.
DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
KW Oxidoreductase.
SQ SEQUENCE 364 AA; 40370 MW; 1B30982E0D5B7E8 CRC64;

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVWFPQAPKDPHWR 323
DB 281 GDVWFPQAPKDPHWR 296
|||||
PRT: 364 AA.
PRELIMINARY;

RESULT 9
ID 000498 . PRELIMINARY; PRT: 364 AA.
AC 000498;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE NAD-dependent formate dehydrogenase.
OS Candida methyllica.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
ON NCBI_TaxID=45353;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96009885; PubMed=7557425;
RX Allen S.J., Holbrook J.J.;
RT "Isolation, sequence and overexpression of the gene encoding NAD-
RT dependent formate dehydrogenase from the methylophilic yeast Candida
RT methyllica.";
RL Gene 162:99-104(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Hall L.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: X81129; CAAS7036.1; -.

```

```

DR HSSP: P33160; 2NAC.
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR Pfam: PF00389; 2-Hacid_DH.1.
DR Pfam: PF02826; 2-Hacid_DH.C.1.
DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.
DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
KW NAD.
SQ SEQUENCE 364 AA; 40344 MW; D83DD1FA9D5EA303 CRC64;

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVWFPQAPKDPHWR 323
DB 281 GDVWFPQAPKDPHWR 296
|||||
PRT: 364 AA.
PRELIMINARY;

RESULT 10
ID 093968 . PRELIMINARY; PRT: 364 AA.
AC 093968;
DT 01-MAY-1999 (TEMBLrel. 10, Created)
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE Formate dehydrogenase (EC 1.2.1.2).
GN FPH.
OS Candida boidinii (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
ON NCBI_TaxID=5477;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCYC 1513;
RX MEDLINE=97369815; PubMed=9226256;
RA Sakai Y., Murdanoto A.P., Konishi T., Iwamatsu A., Kato N.;
RT "Regulation of the formate dehydrogenase gene, FPH1, in the
RT methylophilic yeast Candida boidinii and growth characteristics of an
RT FPH1-disrupted strain on methanol, methylamine, and choline.";
RL J. Bacteriol. 179:4480-4485(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NCYC 1513;
RA Labrou E.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: A1011046; CA09466.2; -.
DR HSSP: P33160; 2NAC.
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR Pfam: PF00389; 2-Hacid_DH.1.
DR Pfam: PF02826; 2-Hacid_DH.C.1.
DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.
DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
KW Oxidoreductase.
SQ SEQUENCE 364 AA; 40266 MW; 34B72787493D7B3F CRC64;

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVWFPQAPKDPHWR 323
DB 281 GDVWFPQAPKDPHWR 296
|||||
PRT: 398 AA.
PRELIMINARY;

RESULT 11
ID 09F7P9 . PRELIMINARY; PRT: 398 AA.
AC 09F7P9;
DT 01-MAR-2001 (TEMBLrel. 16, Created)

```

```
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DE 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DR Predicted NAD-dependent formate dehydrogenase.
OS Uncultured Proteobacterium EBAC31A08.
OC Bacteria; Proteobacteria; environmental samples.
OX NCBI_TaxID=133804;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20446260; PubMed=10988064;
RA Beja O., Aravind L., Koonin E.V., Suzuki M.T., Hadd A., Nguyen L.P.,
RA Jovanovich S.B., Gates C.M., Feldman R.A., Spudich J.L., Spudich E.N.,
RA DeLong E.F.;
RT "Bacterial rhodopsin: evidence for a new type of phototrophy in the
RT sea.";
RL Science 289:1902-1906(2000).
DR EMBL; AF279106; AAG10470.1; -.
DR HSSP; P33160; 2NAD.
DR InterPro; IPR006139; 2-Hacid_DH.
DR InterPro; IPR006140; 2-Hacid_DH_C.
DR Pfam; PF00389; 2-Hacid_DH.1.
DR Pfam; PF02826; 2-Hacid_DH_C.1.
DR PROSITE; PS00065; D_2_HYDROXYACID_DH.1; 1.
DR PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.
DR PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.
SQ SEQUENCE 398 AA; 43992 MW; 7C2EC76FD5723C4C CRC64;

Query Match 3.2%; Score 13; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 GYADVWFPPAP 317
Db 304 GYADVWFPPAP 316
|||||
|||||

RESULT 12
Q08988 ID Q08988 PRELIMINARY; PRT; 236 AA.
AC Q08988:
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE ORF YPL275W.
GN YPL275W.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H., Hebling U.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Duesterhoeft A., Floeth M., Fritz M., Hilbert H., Moestl D.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z73632; CA98013.1; -.
DR HSSP; P33160; 2NAD.
DR InterPro; IPR006140; 2-Hacid_DH_C.
DR Pfam; PF02826; 2-Hacid_DH_C.1.
DR PROSITE; PS00670; D_2_HYDROXYACID_DH_2; 1.
DR PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.
DR PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.
SQ SEQUENCE 236 AA; 26487 MW; 1C8EA068460D166 CRC64;

Query Match 3.0%; Score 12; DB 3; Length 236;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 314 QPAKDPHWRM 325
|||||
|||||

Db 161 QPAKDPHWRM 172

RESULT 13
Q08911 ID Q08911 PRELIMINARY; PRT; 376 AA.
AC Q08911:
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Chromosome XV reading frame ORF YOR388C.
GN FDI1 OR YOR388C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H., Hebling U., Hofmann B.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z75296; CA99720.1; -.
DR HSSP; P33160; 2NAD.
DR SCD; S005915; FDI1.
DR InterPro; IPR006139; 2-Hacid_DH.
DR InterPro; IPR006140; 2-Hacid_DH_C.
DR Pfam; PF00389; 2-Hacid_DH.1.
DR Pfam; PF02826; 2-Hacid_DH.1.
DR PROSITE; PS00670; D_2_HYDROXYACID_DH_2; 1.
DR PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.
DR PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.
SQ SEQUENCE 376 AA; 41714 MW; 67ECD6F9DDC2A02 CRC64;

Query Match 3.0%; Score 12; DB 3; Length 376;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 314 QPAKDPHWRM 325
Db 301 QPAKDPHWRM 312
|||||
|||||

RESULT 14
Q085C4 ID Q085C4 PRELIMINARY; PRT; 152 AA.
AC Q085C4:
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical conserved protein.
GN OB0717.
OS Oceanobacillus thelyensis.
OC Bacteria; Firmicutes; Bacillales; Oceanobacillus.
OX NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=2220767; PubMed=12235376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus thelyensis isolated from the Iheya
RT ridge and its unexpected adaptive capabilities to extreme
RT environments.";
RT Nucleic Acids Res. 30:3927-3935(2002).
DR EMBL; AP004595; BAC12673.1; -.
DR EMBL; AP004595; BAC12673.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 152 AA; 17117 MW; 5C0BC4955C66D1D CRC64;

Query Match 2.7%; Score 11; DB 16; Length 152;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 313 POPAPKDPWR 323
 |||||
 Db 82 POPAPKDPWR 92

RESULT 15
 Q8W520

ID Q8W520 PRELIMINARY; PRT; 199 AA.
 AC Q8W520;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Formate dehydrogenase (Fragment).
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID:4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang H., Bohnert H.;
 RT "Plant one-carbon genes";
 RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF439732; AAC33598.1;
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR Pfam: PF00389; 2-Hacid_DH_C.
 DR Pfam: PF02826; 2-Hacid_DH_C.
 DR PROSITE: PS00665; D_2_HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 199 AA; 21445 MW; DEB9F8791861B6E9 CRC64;

Query Match 2.7%; Score 11; DB 10; Length 199;
 Best Local Similarity 100.0%; Pred. No. 0.013;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVWFPQAPK 318
 |||||
 Db 188 GDVWFPQAPK 198

Search completed: July 24, 2003, 21:01:22
 Job time : 103 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:31:32 ; Search time 27 Seconds

(Without alignments)
628.395 Million cell updates/sec

Title: US-09-996-008b-2

Perfect score: 2135

Sequence: 1 MAKVLCVLYDDPDVGYPKTY.....YSKGNATGSEBAKFKKAV 401

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued_patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCtUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfilltest.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	823	38.5	364	US-09-203-893A-2	Sequence 2, Appl1
2	820	38.4	364	US-09-203-893A-6	Sequence 6, Appl1
3	818	38.3	364	US-08-817-926-51	Sequence 51, Appl1
4	818	38.3	364	US-09-203-893A-32	Sequence 32, Appl1
5	817	38.3	364	US-09-203-893A-4	Sequence 4, Appl1
6	814	38.1	364	US-09-203-893A-16	Sequence 16, Appl1
7	813	38.1	364	US-09-203-893A-14	Sequence 14, Appl1
8	813	38.1	364	US-09-203-893A-18	Sequence 18, Appl1
9	811	38.0	364	US-09-203-893A-22	Sequence 22, Appl1
10	810	37.9	364	US-09-203-893A-20	Sequence 20, Appl1
11	810	37.9	364	US-09-203-893A-24	Sequence 24, Appl1
12	809	37.9	364	US-09-203-893A-12	Sequence 12, Appl1
13	808	37.8	364	US-09-203-893A-8	Sequence 8, Appl1
14	808	37.8	364	US-09-203-893A-10	Sequence 10, Appl1
15	808	37.8	364	US-09-203-893A-28	Sequence 28, Appl1
16	807	37.8	364	US-09-203-893A-26	Sequence 26, Appl1
17	807	37.8	364	US-09-203-893A-30	Sequence 30, Appl1
18	310	14.5	339	US-08-328-352-7424	Sequence 7424, Ap
19	309.5	14.5	345	US-09-222-817-2	Sequence 2, Appl1
20	309.5	14.5	345	US-09-222-786-2	Sequence 2, Appl1
21	309.5	14.5	530	US-09-222-817-12	Sequence 12, Appl1
22	309.5	14.5	530	US-09-222-786-12	Sequence 12, Appl1
23	309	14.5	539	US-09-134-001C-5118	Sequence 5118, Ap
24	307	14.4	530	US-08-222-817-14	Sequence 14, Appl1
25	307	14.4	530	US-08-222-786-14	Sequence 14, Appl1
26	283.5	13.3	375	US-09-252-991A-30745	Sequence 30745, A
27	276	12.9	328	US-08-956-012-1	Sequence 1, Appl1

28	265	12.4	439	1	US-08-553-999B-2	Sequence 2, Appl1
29	263.5	12.3	485	4	US-09-252-991A-25242	Sequence 25242, A
30	258	12.1	410	3	US-08-411-760-14	Sequence 14, Appl1
31	258	12.1	443	4	US-09-252-991A-27545	Sequence 27545, A
32	252	11.8	322	2	US-08-286-819A-2	Sequence 2, Appl1
33	252	11.8	322	3	US-08-980-357-2	Sequence 2, Appl1
34	252	11.8	2291	2	US-08-286-819A-29	Sequence 29, Appl1
35	252	11.8	2291	3	US-08-980-357-29	Sequence 29, Appl1
36	251.5	11.8	1113	4	US-09-252-991A-29215	Sequence 29215, A
37	249	11.7	429	4	US-09-328-352-6282	Sequence 6282, Ap
38	248.5	11.6	335	4	US-09-134-001C-5021	Sequence 5021, Ap
39	246.5	11.5	339	4	US-09-107-532A-7148	Sequence 7148, Ap
40	241	11.3	291	4	US-09-107-532A-5622	Sequence 5622, Ap
41	234.5	11.0	349	4	US-09-134-001C-4884	Sequence 4884, Ap
42	205	9.6	445	4	US-09-252-991A-20277	Sequence 20277, A
43	201	9.4	563	4	US-09-252-991A-24443	Sequence 24443, A
44	194	9.1	326	4	US-09-328-352-4608	Sequence 4608, Ap
45	191	8.9	356	4	US-09-107-532A-6936	Sequence 6936, Ap

ALIGNMENTS

RESULT 1
US-09-203-893A-2
Sequence 2, Application US/09203893A
Patent No. 6242234
GENERAL INFORMATION:
APPLICANT: Kula, Maria-Regina
TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
TITLE OF INVENTION: boldint, new Gene sequences encoding these and use of
TITLE OF INVENTION: the new Formate dehydrogenases
FILE REFERENCE: DEGU 3.0-012
CURRENT APPLICATION NUMBER: US/09/203, 893A
CURRENT FILING DATE: 1998-12-02
PRIOR APPLICATION NUMBER: DE 19753350.7
PRIOR FILING DATE: 1997-12-03
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 364
TYPE: PRT
ORGANISM: Candida boidin1
US-09-203-893A-2

Query Match 38.5%; Score 823; DB 3; Length 364;

Best Local Similarity 45.1%; Pred. No. 3.5e-81;
Matches 171; Conservative 57; Mismatches 113; Indels 38; Gaps 6;

QY	3	KVLGVLYDDPDVGYPKTYARDLPKIDHYPGGQILPFPKALIDFTPGQLGSGVSGELGRP	62
DB	2	KIVLVLD-----AGKHADE-----EKLYGSTENKLGIAN	32
QY	63	YLESNGHLYVTSKDPGDSFEBELVDADVYISQFEPAYLTPERIKAKNLKLTATAG	122
DB	33	WLKQGHLLITTSKDEGSTSLDKHHPADIIITTPRPATITTKERLDKANKLKVVAG	92
QY	123	IGSDHVDLQ--SAIDRNTVAEVTYCNISVAEHVVMILSLVRNYLPDSHEMARKGNNI	180
DB	93	VGSIDHIDYDINGQKGISVLEVTGSNVSAEHHVVMVLVLRNFVPAHQIINHDEW	152
QY	181	ADCSYHVDLEAMVGVTAAGRIGIAVRLAPRD-VILHTYDTRRLRPDSVEKELNLTWN	239
DB	153	AAIKADVDIGKTIATIGAGRIGRIVLERLLPFPKELLYDYQALKEKEKVGARRV	212
QY	240	ATREDMYVCDVVTLMCPLEPTEHEMINDETLLFKRGAYIVNTARGKLCRDVAARALE	299
DB	213	ENIEELVAGADIVVYVNPALHNGTKGLINKELLSKFKAMLVNTARGAICVAEDVAAALE	272
QY	300	SGRLAGTAGDVWPPQAPAKDHPWRTM-----PYNGMPHISGTTLLTAQARYAAGTRETLE	354
DB	273	SGQLRGYGVGDVWFPQAPAKDHPWRDMRKYAGNAMPHYSGTLLDQTRVAEGTKNILE	332


```

RESULT 4
US-09-203-893A-32
; Sequence 32, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boldin1, new Gene sequences encoding these and use of
; TITLE OF INVENTION: the new Formate dehydrogenases
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203.893A
; PRIOR APPLICATION NUMBER: DE 19753350.7
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 32
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boldin1
US-09-203-893A-32

```

```

Query Match      38.3%; Score 818; DB 3; Length 364;
Best Local Similarity 44.9%; Pred. No. 1.2e-80;
Matches 170; Conservative 57; Mismatches 114; Indels 38; Gaps 6;

```

```

QY 3 KVLCLVYDDPVGYPKTYARDLPKIDHYPGQILPPPKAIDFTPGQLGSGVSGELGRP 62
DB 2 KIVLVLD-----AGKHADE-----EKLYGCTENKLGIAN 32
QY 63 YLESGHTLVYTSKDGPDSEFERELVDADVVISQPPRPALTPPERIAKANKLALTAG 122
DB 33 WLKQGHLLTTSKEGTSSELDKHIPDADIIITPPPAITTEKRLDKANKLVLVYAG 92
QY 123 IGSDBVLDQ--SAIDRNVTAEVYTCNSISVAEHVMMILSLVNYLPESHMARKGMI 180
DB 93 VGSIDHLDYINQTKISVLEVTGSNNVSAEHVMTMLVLRNFPANHOIINHMEV 152
QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPD-VLHAYDRHRLPESYEKELNTW 239
DB 153 AAIKADAYDIEGKTATIGAGRIGYRLERLLPKNPKELLYDQALPKAEKVGARV 212
QY 240 ATRDMPVCDVYVTLNCPHETEMINDELKLFKRGATVNTARGLCDRDVAARALE 299
DB 213 ENIELVAQADIVVNPRLAGTGLINKELLSFKKGAMLVNTARGAICVAEDVAAALE 272
QY 300 SGRLAGYAGDVWFQAPAKDHPWRTM-----PYNGMTPHISGTTLTAQARYAGTRETLE 354
DB 273 SGQLRGIGGDVWFQAPAKDHPWRTM-----PYNGMTPHISGTTLTAQARYAGTRETLE 332
QY 355 CFPEGR-PIRDEYLIYOGG 372
DB 333 SFTGKFDYRPDITLLNG 351

```

```

RESULT 5
US-09-203-893A-4
; Sequence 4, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boldin1, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203.893A
; PRIOR APPLICATION NUMBER: DE 19753350.7
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin Ver. 2.1

```

```

; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boldin1
US-09-203-893A-4

```

```

Query Match      38.3%; Score 817; DB 3; Length 364;
Best Local Similarity 44.9%; Pred. No. 1.6e-80;
Matches 170; Conservative 57; Mismatches 114; Indels 38; Gaps 6;

```

```

QY 3 KVLCLVYDDPVGYPKTYARDLPKIDHYPGQILPPPKAIDFTPGQLGSGVSGELGRP 62
DB 2 KIVLVLD-----AGKHADE-----EKLYGCTENKLGIAN 32
QY 63 YLESGHTLVYTSKDGPDSEFERELVDADVVISQPPRPALTPPERIAKANKLALTAG 122
DB 33 WLKQGHLLTTSKEGTSSELDKHIPDADIIITPPPAITTEKRLDKANKLVLVYAG 92
QY 123 IGSDBVLDQ--SAIDRNVTAEVYTCNSISVAEHVMMILSLVNYLPESHMARKGMI 180
DB 93 VGSIDHLDYINQTKISVLEVTGSNNVSAEHVMTMLVLRNFPANHOIINHMEV 152
QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPD-VLHAYDRHRLPESYEKELNTW 239
DB 153 AAIKADAYDIEGKTATIGAGRIGYRLERLLPKNPKELLYDQALPKAEKVGARV 212
QY 240 ATRDMPVCDVYVTLNCPHETEMINDELKLFKRGATVNTARGLCDRDVAARALE 299
DB 213 ENIELVAQADIVVNPRLAGTGLINKELLSFKKGAMLVNTARGAICVAEDVAAALE 272
QY 300 SGRLAGYAGDVWFQAPAKDHPWRTM-----PYNGMTPHISGTTLTAQARYAGTRETLE 354
DB 273 SGQLRGIGGDVWFQAPAKDHPWRTM-----PYNGMTPHISGTTLTAQARYAGTRETLE 332
QY 355 CFPEGR-PIRDEYLIYOGG 372
DB 333 SFTGKFDYRPDITLLNG 351

```

```

RESULT 6
US-09-203-893A-16
; Sequence 16, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boldin1, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203.893A
; PRIOR APPLICATION NUMBER: DE 19753350.7
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 16
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boldin1
US-09-203-893A-16

```

```

Query Match      38.1%; Score 814; DB 3; Length 364;
Best Local Similarity 44.9%; Pred. No. 3.4e-80;
Matches 170; Conservative 57; Mismatches 114; Indels 38; Gaps 6;

```

```

QY 3 KVLCLVYDDPVGYPKTYARDLPKIDHYPGQILPPPKAIDFTPGQLGSGVSGELGRP 62
DB 2 KIVLVLD-----AGKHADE-----EKLYGCTENKLGIAN 32
QY 63 YLESGHTLVYTSKDGPDSEFERELVDADVVISQPPRPALTPPERIAKANKLALTAG 122
DB 33 WLKQGHLLTTSKEGTSSELDKHIPDADIIITPPPAITTEKRLDKANKLVLVYAG 92

```


SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 22
LENGTH: 364
TYPE: PRT
ORGANISM: Candida boidinii
US-09-203-893A-22

Query Match 38.0%; Score 811; DB 3; Length 364;
Best Local Similarity 44.6%; Pred. No. 7,2e-80;
Matches 169; Conservative 58; Mismatches 114; Indels 38; Gaps 6;

QY 3 KVLCLYDDPVGYPKTYARDLPKIDHYPGQILPTPKAIDFTPGQLGVSSELGRP 62
DB 2 KIVLVLD-----AGKHADE-----EKLYGATEMKLGIAN 32
QY 63 YLENGHTLVVTSKDGPDVSFERELVDADVVISOPFPAYLTPEIRAKANKLALTAG 122
DB 33 WLKOGHELITTSKEGETSELDKHIPDADIIITPPHAYITKERLDKANKLVLVAVG 92
QY 123 IGSDDVDLQ--SAIDRNVTAEVTCNSISVAEHVMMILSLVRNYLPSEHEMARGGNI 180
DB 93 VGSDDHIDLDYINQGGKISVLEVTGSNVSAEHVMTLVLRNFPVPAHQIINHDEY 152
QY 181 ADCVSHAYDLEAMHGVTAAGRIGLAVLRRLAPD-VHLHYTDHRRRLPESVEKELNTW 239
DB 153 AAIKADAVDIEGKTATIGAGRIGYRVLRLPFPKELLYDYOALPKAEKVGARV 212
QY 240 ATREDMPVCDDVYVTLNCPHETEMINDELKLFKRGATVNTARGLCDRAVAAALE 299
DB 213 ENIELVAQADIVTVNAPLHAGTGLINKELLSFKKGAWLVNARGAIVAAEDVAAALE 272
QY 300 SGRLAGYAGDVWFPQAPAKDHPWRTM-----PYNGMPHISGTTLTAQARYAAGRTRE 354
DB 273 SGOLRGYGDDWFPQAPAKDHPWDMRNKRYGAGNAMPHYSGTLLDQOTRIABEGTKNLE 332
QY 355 CFPEGR-PIRDEYLVGG 372
DB 333 SFTGKFDYRPQDITLLNG 351

RESULT 10
US-09-203-893A-20
Sequence 20, Application US/09203893A
Patent No. 6242234
GENERAL INFORMATION:
APPLICANT: Kula, Maria-Regina
TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
TITLE OF INVENTION: boidinii, new gene sequences encoding these and use of
FILE REFERENCE: DECU 3.0-012
CURRENT APPLICATION NUMBER: US/09/203,893A
PRIOR FILING DATE: 1998-12-02
PRIOR APPLICATION NUMBER: DE 19753350.7
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 20
LENGTH: 364
TYPE: PRT
ORGANISM: Candida boidinii
US-09-203-893A-20

Query Match 37.9%; Score 810; DB 3; Length 364;
Best Local Similarity 44.6%; Pred. No. 9,3e-80;
Matches 169; Conservative 58; Mismatches 114; Indels 38; Gaps 6;

QY 3 KVLCLYDDPVGYPKTYARDLPKIDHYPGQILPTPKAIDFTPGQLGVSSELGRP 62
DB 2 KIVLVLD-----AGKHADE-----EKLYGATEMKLGIAN 32
QY 63 YLENGHTLVVTSKDGPDVSFERELVDADVVISOPFPAYLTPEIRAKANKLALTAG 122
DB 33 WLKOGHELITTSKEGETSELDKHIPDADIIITPPHAYITKERLDKANKLVLVAVG 92

QY 123 IGSDDVDLQ--SAIDRNVTAEVTCNSISVAEHVMMILSLVRNYLPSEHEMARGGNI 180
DB 93 VGSDDHIDLDYINQGGKISVLEVTGSNVSAEHVMTLVLRNFPVPAHQIINHDEY 152
QY 181 ADCVSHAYDLEAMHGVTAAGRIGLAVLRRLAPD-VHLHYTDHRRRLPESVEKELNTW 239
DB 153 AAIKADAVDIEGKTATIGAGRIGYRVLRLPFPKELLYDYOALPKAEKVGARV 212
QY 240 ATREDMPVCDDVYVTLNCPHETEMINDELKLFKRGATVNTARGLCDRAVAAALE 299
DB 213 ENIELVAQADIVTVNAPLHAGTGLINKELLSFKKGAWLVNARGAIVAAEDVAAALE 272
QY 300 SGRLAGYAGDVWFPQAPAKDHPWRTM-----PYNGMPHISGTTLTAQARYAAGRTRE 354
DB 273 SGOLRGYGDDWFPQAPAKDHPWDMRNKRYGAGNAMPHYSGTLLDQOTRIABEGTKNLE 332
QY 355 CFPEGR-PIRDEYLVGG 372
DB 333 SFTGKFDYRPQDITLLNG 351

RESULT 11
US-09-203-893A-24
Sequence 24, Application US/09203893A
Patent No. 6242234
GENERAL INFORMATION:
APPLICANT: Kula, Maria-Regina
TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
TITLE OF INVENTION: boidinii, new gene sequences encoding these and use of
FILE REFERENCE: DECU 3.0-012
CURRENT APPLICATION NUMBER: US/09/203,893A
PRIOR FILING DATE: 1998-12-02
PRIOR APPLICATION NUMBER: DE 19753350.7
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 24
LENGTH: 364
TYPE: PRT
ORGANISM: Candida boidinii
US-09-203-893A-24

Query Match 37.9%; Score 810; DB 3; Length 364;
Best Local Similarity 44.6%; Pred. No. 9,3e-80;
Matches 169; Conservative 58; Mismatches 114; Indels 38; Gaps 6;

QY 3 KVLCLYDDPVGYPKTYARDLPKIDHYPGQILPTPKAIDFTPGQLGVSSELGRP 62
DB 2 KIVLVLD-----AGKHADE-----EKLYGATEMKLGIAN 32
QY 63 YLENGHTLVVTSKDGPDVSFERELVDADVVISOPFPAYLTPEIRAKANKLALTAG 122
DB 33 WLKOGHELITTSKEGETSELDKHIPDADIIITPPHAYITKERLDKANKLVLVAVG 92
QY 123 IGSDDVDLQ--SAIDRNVTAEVTCNSISVAEHVMMILSLVRNYLPSEHEMARGGNI 180
DB 93 VGSDDHIDLDYINQGGKISVLEVTGSNVSAEHVMTLVLRNFPVPAHQIINHDEY 152
QY 181 ADCVSHAYDLEAMHGVTAAGRIGLAVLRRLAPD-VHLHYTDHRRRLPESVEKELNTW 239
DB 153 AAIKADAVDIEGKTATIGAGRIGYRVLRLPFPKELLYDYOALPKAEKVGARV 212
QY 240 ATREDMPVCDDVYVTLNCPHETEMINDELKLFKRGATVNTARGLCDRAVAAALE 299
DB 213 ENIELVAQADIVTVNAPLHAGTGLINKELLSFKKGAWLVNARGAIVAAEDVAAALE 272
QY 300 SGRLAGYAGDVWFPQAPAKDHPWRTM-----PYNGMPHISGTTLTAQARYAAGRTRE 354
DB 273 SGOLRGYGDDWFPQAPAKDHPWDMRNKRYGAGNAMPHYSGTLLDQOTRIABEGTKNLE 332
QY 355 CFPEGR-PIRDEYLVGG 372

Db 333 SFTGKFDYRPDIILLNG 351

RESULT 12
US-09-203-893A-12
Sequence 12, Application US/09203893A
Patent No. 6242234

GENERAL INFORMATION:
APPLICANT: Kula, Maria-Regina
TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
TITLE OF INVENTION: boldin1, new Gene sequences encoding these and use of
FILE REFERENCE: DEGU 3.0-012
CURRENT APPLICATION NUMBER: US/09/203,893A
PRIORITY FILING DATE: 1998-12-02
PRIOR APPLICATION NUMBER: DE 19753350.7
PRIOR FILING DATE: 1997-12-03
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 12
LENGTH: 364
TYPE: PRF
ORGANISM: Candida boldin1
US-09-203-893A-12

Query Match 37.8%; Score 809; DB 3; Length 364;
Best Local Similarity 44.6%; Pred. No. 1.2e-79;
Matches 169; Conservative 57; Mismatches 115; Indels 38; Gaps 6;

QY 3 KVLCLYDDPVDGPKTYARDLPRKIDHFGGILPTPKAIDFTPGQLGSVSGELGRLP 62
Db 2 KIVLVLYD-----AGKHADE-----EKLYGCTENKLGIAN 32
QY YLESNGHTLVVTSKOPDSVFERELVDADVISOPEWPAVLPERRAKKNKLALTAG 122
Db 33 WLDQGHLLTTSDKESETSELDKHIPDADIIITTPHPAVITRKERLDKKNKLVLVAG 92
QY 123 IGSDDHVDLQ--SAIDRNVTAEVYTCNSISVAEHVMMILSVRNYLPSHEMARCKGWN 180
Db 93 VGSDDHLDLYINOTGKKISVLEVYTGSNVSVAEHVMTMLVLRNFPAPHAEOIINHMEV 152
QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVRLRLAPD-VLHLYTDRHRLPESVEKELNTW 239
Db 153 AAIKADAYDIEGKTATIGAGRIGYRLERLPLFPNRELLYDYQALPKAEAEKVGARV 212
QY 240 ATREDMYPVCDVYTLNCPRLPTEHMTINDETLKLFKGAIVMTARGKLCDDRAVARALE 299
Db 213 ENIEELVAQDIYTVNAPRLAGTKGLINKELLSKFKGAWLVNTARGAIVAEVAAALE 272
QY 300 SGRLAGYAGDVPFPPAPKDPKRTM----PYNGMTPHISGTTLTAAQARYAAGTREILE 354
Db 273 SGOLRGYGDVWFPQAPKDPKRTM----PYNGMTPHISGTTLTAAQARYAAGTREILE 332
QY 355 CFEEGR-PIRDEYLIYOGG 372
Db 333 SFTGKFDYRPDIILLNG 351

RESULT 13
US-09-203-893A-8
Sequence 8, Application US/09203893A
Patent No. 6242234

GENERAL INFORMATION:
APPLICANT: Kula, Maria-Regina
TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
TITLE OF INVENTION: boldin1, new Gene sequences encoding these and use of
FILE REFERENCE: DEGU 3.0-012
CURRENT APPLICATION NUMBER: US/09/203,893A
PRIORITY FILING DATE: 1998-12-02
PRIOR APPLICATION NUMBER: DE 19753350.7
PRIOR FILING DATE: 1997-12-03

NUMBER OF SEQ ID NOS: 37
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 8
LENGTH: 364
TYPE: PRF
ORGANISM: Candida boldin1
US-09-203-893A-8

Query Match 37.8%; Score 808; DB 3; Length 364;
Best Local Similarity 44.6%; Pred. No. 1.5e-79;
Matches 169; Conservative 57; Mismatches 115; Indels 38; Gaps 6;

QY 3 KVLCLYDDPVDGPKTYARDLPRKIDHFGGILPTPKAIDFTPGQLGSVSGELGRLP 62
Db 2 KIVLVLYD-----AGKHADE-----EKLYGCTENKLGIAN 32
QY YLESNGHTLVVTSKOPDSVFERELVDADVISOPEWPAVLPERRAKKNKLALTAG 122
Db 33 WLDQGHLLTTSDKESETSELDKHIPDADIIITTPHPAVITRKERLDKKNKLVLVAG 92
QY 123 IGSDDHVDLQ--SAIDRNVTAEVYTCNSISVAEHVMMILSVRNYLPSHEMARCKGWN 180
Db 93 VGSDDHLDLYINOTGKKISVLEVYTGSNVSVAEHVMTMLVLRNFPAPHAEOIINHMEV 152
QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVRLRLAPD-VLHLYTDRHRLPESVEKELNTW 239
Db 153 AAIKADAYDIEGKTATIGAGRIGYRLERLPLFPNRELLYDYQALPKAEAEKVGARV 212
QY 240 ATREDMYPVCDVYTLNCPRLPTEHMTINDETLKLFKGAIVMTARGKLCDDRAVARALE 299
Db 213 ENIEELVAQDIYTVNAPRLAGTKGLINKELLSKFKGAWLVNTARGAIVAEVAAALE 272
QY 300 SGRLAGYAGDVPFPPAPKDPKRTM----PYNGMTPHISGTTLTAAQARYAAGTREILE 354
Db 273 SGOLRGYGDVWFPQAPKDPKRTM----PYNGMTPHISGTTLTAAQARYAAGTREILE 332
QY 355 CFEEGR-PIRDEYLIYOGG 372
Db 333 SFTGKFDYRPDIILLNG 351

RESULT 14
US-09-203-893A-10
Sequence 10, Application US/09203893A
Patent No. 6242234

GENERAL INFORMATION:
APPLICANT: Kula, Maria-Regina
TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
TITLE OF INVENTION: boldin1, new Gene sequences encoding these and use of
FILE REFERENCE: DEGU 3.0-012
CURRENT APPLICATION NUMBER: US/09/203,893A
PRIORITY FILING DATE: 1998-12-02
PRIOR APPLICATION NUMBER: DE 19753350.7
PRIOR FILING DATE: 1997-12-03
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 10
LENGTH: 364
TYPE: PRF
ORGANISM: Candida boldin1
US-09-203-893A-10

Query Match 37.8%; Score 808; DB 3; Length 364;
Best Local Similarity 44.6%; Pred. No. 1.5e-79;
Matches 169; Conservative 57; Mismatches 115; Indels 38; Gaps 6;

QY 3 KVLCLYDDPVDGPKTYARDLPRKIDHFGGILPTPKAIDFTPGQLGSVSGELGRLP 62
Db 2 KIVLVLYD-----AGKHADE-----EKLYGCTENKLGIAN 32
QY YLESNGHTLVVTSKOPDSVFERELVDADVISOPEWPAVLPERRAKKNKLALTAG 122
Db 33 WLDQGHLLTTSDKESETSELDKHIPDADIIITTPHPAVITRKERLDKKNKLVLVAG 92

Db 33 WLKDGHELIITTSDEKGETSELDKHIIPDADIIITTFPHPAYITKERLDKAKNLKLVVAG 92
QY 123 IGSDBVDLQ--SALDRNTVAEVTYCNLSVAEHVMMILSLVRYLPSHEARKGNMI 180
Db 93 VGSDBHIDIDYINQIGKKTISVEVTGSNVVAEHHVMTMLVYRNFVPAHEQIINHDEV 152
QY 181 ADCVSHAYDLBAMHGTVAAGRIGLAVLRRLAPFD-VHLHYTDHRRLPESVEKEINLTMH 239
Db 153 AAIKADAYDIEGKTATIGAGRIGRYVLERLLPFNPKEKLYDYDQALPKEAEKVGARV 212
QY 240 ATREDMYPVCVYTLNCLPHETEMINDETLKFKGAYIVNTARGKLCRDVAVARALE 299
Db 213 ENIELVAQADIVYFNAPLHAGTKGLINKELSKFKKGMVNTARGAIVVAEDVAAALE 272
QY 300 SGRLAGYAGDVWFPQAPKDPHWRM-----PYNGMTPHISGTTLTAQARYAGTRETLE 354
Db 273 SGQLRGYGDVWFPQAPKDPHWRMKNKYGAGNMTPHYSGTTLDAQTRYAEGTKNILE 332
QY 355 CFEEGR-PIRDEYLIWOGG 372
Db 333 SFTGKFDYRPDIILLNG 351

RESULT 15
US-09-203-893A-28
; Sequence 28, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boidin1, new Gene sequences encoding these and use of
; TITLE OF INVENTION: the new Formate dehydrogenases
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; CURRENT FILING DATE: 1996-12-02
; PRIOR APPLICATION NUMBER: DE 19753350.7
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boidin1
US-09-203-893A-28

Query Match 37.8%; Score 808; DB 3; Length 364;
Best Local Similarity 44.6%; Pred. No. 1.5e-79;
Matches 169; Conservative 57; Mismatches 115; Indels 38; Gaps 6;

QY 3 KVLCLVYDDPYDGYRKYARDDLPRKIDHYPGQILPTPKAIDFTPGQLLSVSGELGRP 62
Db 2 KIVLVLYD-----AGKHADE-----EKLXGVTENKLGIAN 32
QY 63 YLENGHTLVYTSKDGPDVSFERELVDADVVISQFPWPAYLTPERIAKAKNLKLTATAG 122
Db 33 WLKDGHELIITTSDEKGETSELDKHIIPDADIIITTFPHPAYITKERLDKAKNLKLVVAG 92
QY 123 IGSDBVDLQ--SALDRNTVAEVTYCNLSVAEHVMMILSLVRYLPSHEARKGNMI 180
Db 93 VGSDBHIDIDYINQIGKKTISVEVTGSNVVAEHHVMTMLVYRNFVPAHEQIINHDEV 152
QY 181 ADCVSHAYDLBAMHGTVAAGRIGLAVLRRLAPFD-VHLHYTDHRRLPESVEKEINLTMH 239
Db 153 AAIKADAYDIEGKTATIGAGRIGRYVLERLLPFNPKEKLYDYDQALPKEAEKVGARV 212
QY 240 ATREDMYPVCVYTLNCLPHETEMINDETLKFKGAYIVNTARGKLCRDVAVARALE 299
Db 213 ENIELVAQADIVYFNAPLHAGTKGLINKELSKFKKGMVNTARGAIVVAEDVAAALE 272
QY 300 SGRLAGYAGDVWFPQAPKDPHWRM-----PYNGMTPHISGTTLTAQARYAGTRETLE 354
Db 273 SGQLRGYGDVWFPQAPKDPHWRMKNKYGAGNMTPHYSGTTLDAQTRYAEGTKNILE 332

QY 355 CFEEGR-PIRDEYLIWOGG 372
Db 333 SFTGKFDYRPDIILLNG 351

Search completed: July 24, 2003, 20:40:28
Job time : 30 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:39:02 ; Search time 29 Seconds

(without alignments)
1642.161 Million cell updates/sec

Title: US-09-996-008B-2

Perfect score: 2135
Sequence: 1 MAKVLCVLYDDPVVDGPKTY.....YSKGNATGSGSEAKFKKAV 401

Scoring table:

BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/2/pubpaa/PC1_NEW_PUB.pep:*

3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*

7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*

10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*

11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/2/pubpaa/US09C_NEW_PUB.pep:*

13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*

14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*

17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1428.5	66.9	387	US-10-156-761-9479	Sequence 9479, App
2	838.5	39.3	376	US-10-224-567-1	Sequence 1, Appl
3	781	36.6	343	US-09-815-242-12981	Sequence 12981, A
4	781	36.6	343	US-09-815-242-13150	Sequence 13150, A
5	780.5	36.6	335	US-09-815-242-5830	Sequence 5830, Ap
6	340.5	15.9	325	US-10-156-761-15009	Sequence 15009, A
7	310.5	14.5	530	US-09-738-626-4915	Sequence 4915, Ap
8	309.5	14.5	530	US-10-195-548-12	Sequence 12, Appl
9	307	14.4	530	US-10-195-548-14	Sequence 14, Appl
10	297.5	13.9	529	US-10-156-761-10267	Sequence 10267, A
11	297	13.9	410	US-09-815-242-11028	Sequence 11028, A
12	277	13.0	331	US-09-815-242-11761	Sequence 11761, A
13	275.5	12.9	445	US-10-177-293-92	Sequence 92, Appl
14	264	12.4	985	US-10-205-823-76	Sequence 76, Appl

16	264	12.4	985	15	US-10-177-293-94	Sequence 94, Appl
17	258	12.1	410	9	US-09-815-242-10291	Sequence 10291, A
18	256	12.0	524	9	US-09-815-242-11557	Sequence 11557, A
19	254	11.9	427	9	US-09-815-242-13743	Sequence 13743, A
20	244.5	11.5	329	9	US-09-815-242-11811	Sequence 11811, A
21	234.5	11.0	335	15	US-10-156-761-9627	Sequence 9627, Ap
22	224.5	10.5	304	9	US-09-738-626-6082	Sequence 6082, Ap
23	217.5	10.2	333	9	US-09-815-242-10900	Sequence 10900, A
24	210.5	9.9	336	10	US-09-925-300-1408	Sequence 1408, Ap
25	210	9.8	325	15	US-10-122-466A-17	Sequence 17, Appl
26	206.5	9.7	315	10	US-09-922-501-14	Sequence 14, Appl
27	203.5	9.5	313	15	US-10-156-761-10272	Sequence 10272, A
28	195.5	9.2	319	15	US-10-156-761-12767	Sequence 12767, A
29	192	9.0	304	10	US-09-738-626-3580	Sequence 3580, Ap
30	126	5.9	171	9	US-09-939-980-307	Sequence 307, App
31	101.5	4.8	432	15	US-10-156-761-14180	Sequence 14180, A
32	95	4.4	444	15	US-10-156-761-13239	Sequence 13239, A
33	94	4.4	1144	15	US-10-156-761-7801	Sequence 7801, Ap
34	92	4.3	398	15	US-10-156-761-13771	Sequence 13711, A
35	91	4.3	327	11	US-09-284-320-16	Sequence 16, Appl
36	91	4.3	915	15	US-10-156-761-10050	Sequence 10050, A
37	90.5	4.2	566	9	US-09-815-242-13371	Sequence 13371, A
38	89.5	4.2	71	9	US-09-864-761-46245	Sequence 46245, A
39	87	4.1	792	9	US-09-815-242-11966	Sequence 11966, A
40	87	4.1	792	15	US-10-127-032-165	Sequence 165, App
41	87	4.1	1139	15	US-10-156-761-10856	Sequence 10856, A
42	87	4.1	1426	9	US-09-912-020-340	Sequence 340, App
43	86.5	4.1	3970	15	US-10-156-761-10429	Sequence 10429, A
44	86	4.0	716	15	US-10-156-761-10760	Sequence 10760, A
45	85.5	4.0	415	9	US-09-815-242-11084	Sequence 11084, A

ALIGNMENTS

RESULT 1	US-10-156-761-9479	Application US/10156761
Sequence 9479, Appl	US-10-156-761-9479	
Publication No. US20030119018A1		
GENERAL INFORMATION:		
APPLICANT: OMURA, SATOSHI		
APPLICANT: IKEDA, HARUO		
APPLICANT: ISHIKAWA, JUN		
APPLICANT: HORIKAWA, HIROSHI		
APPLICANT: SHIDA, TADAYOSHI		
APPLICANT: SAKAKI, YOSHIYUKI		
APPLICANT: HATTORI, MASAHIRA		
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES		
FILE REFERENCE: 249-262		
CURRENT APPLICATION NUMBER: US/10/156,761		
PRIOR FILING DATE: 2002-05-29		
PRIOR APPLICATION NUMBER: JP 2001-204089		
PRIOR FILING DATE: 2001-05-30		
PRIOR APPLICATION NUMBER: JP 2001-272697		
PRIOR FILING DATE: 2001-08-02		
NUMBER OF SEQ ID NOS: 15109		
SEQ ID NO 9479		
LENGTH: 387		
TYPE: PRT		
ORGANISM: Streptomyces avermitilis		
US-10-156-761-9479		
Query Match	66.9%	Score 1428.5; DB 15; Length 387;
Best Local Similarity	69.5%	Pred. No. 5e-140;
Matches 267; Conservative	42;	Mismatches 74; Indels 1; Gaps 1;
QY	1	MAKVLCVLYDDPVVDGPKTYARDLPEKIDHVGGLPDKPAIDFTPOLGSGVSGEL 60
DB	1	MAKVLCVLYDDPVVDGPKTYARDLPEKIDHVGGLPDKPAIDFTPOLGSGVSGEL 60
QY	61	RPYLESNCHTLVTVSDKDGDPVSEFERLVADVTVISQFPWATLPERIAKKNLAL 120
DB	61	RPFLEDRDITVTVSDKDGDPVSEFERLVADVTVISQFPWATLPERIAKKNLAL 120

QY 121 AGIGSDHVDLSAIDRNVTVAEYTCNSISVAEHVMMILSLVNTLPSEH-ARKGWN 179
122 AGIGSDHVDLSAIDRNVTVAEYTCNSISVAEHVMMILSLVNTLPSEH-ARKGWN 180
QY 180 IADCVSHADLEMHVGTAAAGRIGLAVLRRLAPFDVHLHYDTRHRLPESVEKELNITWH 239
181 IADSVSAVDLEGMVDVGLSGRIGAVLRRLAPFDVHLHYDTRHRLPESVEKELNITWH 240
QY 240 ATREDVYPCVDVYTLNCPLEPTEHMINDETLLFKRGAVYMTARGKLCDDRAVVALE 239
241 PDARSLASSVDVSTIRPLPQTQNFDDLLGAMKRGSTIVTARLIVDRDAVVALN 300
QY 300 SGRLAGAGDVWPEOPAPKPMRTMPYNGMTPHISGTTTLTAQARYAAGTRELCEPFE 359
301 SGLAGYAGDVWYPPPPPHRTMPYEMTPHVSSTLSAQAARYAAGTRELCEPFE 360
QY 360 RPIRDEYLYOGGALAGTGAHYS 383
361 RPIRPEYLYDGGGAGTGAHYS 384

RESULT 2

US-10-224-567-1
Sequence 1, Application US/10224567
Publication No. US2003008763A1
GENERAL INFORMATION:
APPLICANT: Kobayashi, Akio
APPLICANT: Fukusaki, Eiichihiro
APPLICANT: Isogai, Akira
TITLE OF INVENTION: Method to Promote Growth of a Plant
FILE REFERENCE: 026350-077
CURRENT APPLICATION NUMBER: US/10/224,567
CURRENT FILING DATE: 2002-08-21
PRIOR APPLICATION NUMBER: US 09/517,427
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: JP 11-56,776
PRIOR FILING DATE: 1999-03-04
PRIOR APPLICATION NUMBER: JP 2000-36,153
PRIOR FILING DATE: 2000-02-15
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 376
TYPE: PRT
ORGANISM: Rice formate dehydrogenase
FEATURE:
NAME/KEY: VARIANT
LOCATION: 287
OTHER INFORMATION: Xaa - Any Amino Acid
US-10-224-567-1

Query Match 39.3%; Score 838.5; DB 15; Length 376;
Best Local Similarity 49.3%; Pred. No. 1.4e-78;
Matches 170; Conservative 53; Mismatches 115; Indels 7; Gaps 2;

QY 31 YPGGQLPPPKAIDTPGQLGSGVSELGLRPLYSNGHTLVTSKDGPDVFEERLYD 90
35 YKGEY-----ADKNP-NFVCGVEGALGIREWLESKHNYITVDKEGLNSELEKHTED 87
QY 91 ADVVIOQPPRALITREIRAKAKNLALTAGISGSHVDLSAIDRNVTVAEYTCNSIS 150
88 MHVLTTPPHPAVSAERIKAKNLLELLTAGISGSHIDLPAAAGALVAAETGSGTWS 147
QY 151 VAHVMMILSLVNTLPSEHARKGWNIAIDCVSHADLEMHVGTAAAGRIGLAVLR 210
148 VADELMLRILILRLNLPGLQGVHGEWVAGIATYATIDEGTGTGAGRIGLILQR 207
QY 211 LAFDVHLHYDTRHRLPESVEKELNITWHATREDMTVPCDVVTLNCPLEPTEHMINDET 270
208 LKPFNCNLHYDRILKIDPELEKEIGAKYEEDLDAMLPCDVIYIINTPLETEKGTGMNKR 267
QY 271 LKFKRGAVYMTARGKLCDDRAVVALESGRLAGYAGDVWYPPAPKPMRTMPYNGM 330

DB 268 IAKMKGVILVNNARGAIMTQAVADACSSGQVAGGVDWFPQAPKGPMBRYMNHAM 327
QY 331 TPHTSGTTLTAQARYAAGTRELCEPFEGRINDEYLYOGGALA 375
DB 328 TPHTSGTTLTAQARYAAGTRELCEPFEGRINDEYLYOGGALA 372

RESULT 3

US-09-815-242-12981
Sequence 12981, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlson, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.01A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12981
LENGTH: 343
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12981

Query Match 36.6%; Score 781; DB 9; Length 343;
Best Local Similarity 45.6%; Pred. No. 1.2e-72;
Matches 159; Conservative 64; Mismatches 104; Indels 22; Gaps 6;

QY 40 PKAIDTPGQLGSGVSELGLRPLYSNGHTLVTSKDGPDVFEERLYDVAIVISOPF 99
DB 10 PEAVGEOENLWTKKA-IGLKTFLBERGHEFTLLD-NQED-LDNHLPDMVITISAPF 65
QY 100 WPAVLTPEIRAKAKNLALTAGISGSHVDLSAIDRNVTVAEYTCNSISVAEHVMMIL 159
66 YPAVMTREIRERKAPNLKATITAGVSDHVDLAASEHINIGVETSNTVSAEHAVMOL 125
QY 160 LSLVNTLPSEHARKGWNIAIDCVSHADLEMHVGTAAAGRIGLAVLRRLAPFDVHL 219
126 LILIRVTECHROSVGEWNLQOVGNHAEHQHTIGFGFRIQGLVVERLAPFVNTIQ 185
QY 220 YTD-----RHRLESVEKELNITWHATREDMTVPCDVVTLNCPLEPTEHMINDETLLK 274
186 HDPINQODKLSKEV-----SPDELVSTSAITTHADLPETDNLFDKVDLSRM 235
QY 275 KRGAVYMTARGKLCDDRAVVALESGRLAGYAGDVWFPQAPKPMRTMPYNGMTPH 334
DB 236 KHSYLVNTARGIYVRDALVEALASEHIGYAGDVWYPPAPADHPMTMPNANTVHY 295
QY 335 SGTTLTAQARYAAGTRELCEPFEGRINDEYLYOGGALAGTGAHYS 383

Db 296 SGMTELEOKRIEDGVKDLERFFNHPEPQDDIIVASGRIA---SKSYT 341

RESULT 4

US-09-815-242-13150

; Sequence 13150, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FASTSEQ for Windows Version 4.0

; SEQ ID NO 13150

; LENGTH: 343

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-815-242-13150

Query Match 36.6%; Score 781; DB 9; Length 343;
Best Local Similarity 45.6%; Pred. No. 1.2e-72;
Matches 159; Conservative 64; Mismatches 104; Indels 22; Gaps 6;

40 PKAIDFTPGQLLSVSGELGRPLYESNGHTLVTSKDGDPDSVFERELVADVYISOPF 99
10 PEAVEGQENQNLNTRKA-LGKLTFLERGHETIAD-NGED--LDKHLPMDDVIISAPF 65
100 WPAVTPERIRAKAKNLKLTATAGISGSHVDLSAIDRNVTVAEVTCNSISVAEHWMMI 159
66 YPAVTRERIEREKAAPKLKLTATAGISGSHVDLSAASEHNIGVEYTGSMTVSAEHAAMD 125
160 LSLVRNLYPSHEMARKGGMNIADCVSHAYDLEAMHVGTVAGRGIGLAVLRRLAPEDVILH 219
126 LILLRNVEEGHRSQVSEGMNLSQVGNHAEHLQHTKIGLFGRGIGLVAERLAPFNVTIQ 185
220 YTD-----RHRLPSEVEKELNLTWHTATREDMTYPCVDVYTLNCPHPEEHMINDETLKF 274
186 HYDPIINQDHLKSFV-----SDELVSTSDATITIAAPLPEIDNLFDDKVLSRM 235
275 KRGAIVYNTARGKLCDDRAVAARALESGRLAGAGVWPPAPAKDHPRTMPYNGMTPHI 334
236 KKHSHVYNTARGKIVNRPDALVEALASEHLQGYAGDVWYPPAPADHPRTMPRNMATVHY 295
335 SGTTLTAQARYAAGTRETILECFEGRPIRDEYLYVOGALAGTGAHSY 383
296 SGMTELEOKRIEDGVKDLERFFNHPEPQDDIIVASGRIA---SKSYT 341

RESULT 5

US-09-815-242-5830

; Sequence 5830, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FASTSEQ for Windows Version 4.0

; SEQ ID NO 5830

; LENGTH: 335

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-815-242-5830

Query Match 36.6%; Score 780.5; DB 9; Length 335;
Best Local Similarity 46.0%; Pred. No. 1.3e-72;
Matches 157; Conservative 62; Mismatches 103; Indels 19; Gaps 5;

40 PKAIDFTPGQLLSVSGELGRPLYESNGHTLVTSKDGDPDSVFERELVADVYISOPF 99
7 PEAVEGQENQNLNTRKA-LGKLTFLERGHETIAD-NGED--LDKHLPMDDVIISAPF 62
100 WPAVTPERIRAKAKNLKLTATAGISGSHVDLSAIDRNVTVAEVTCNSISVAEHWMMI 159
63 YPAVTRERIEREKAAPKLKLTATAGISGSHVDLSAASEHNIGVEYTGSMTVSAEHAAMD 122
160 LSLVRNLYPSHEMARKGGMNIADCVSHAYDLEAMHVGTVAGRGIGLAVLRRLAPEDVILH 219
123 LILLRNVEEGHRSQVSEGMNLSQVGNHAEHLQHTKIGLFGRGIGLVAERLAPFNVTIQ 182
220 YTD-----RHRLPSEVEKELNLTWHTATREDMTYPCVDVYTLNCPHPEEHMINDETLKF 274
183 HYDPIINQDHLKSFV-----SDELVSTSDATITIAAPLPEIDNLFDDKVLSRM 232
275 KRGAIVYNTARGKLCDDRAVAARALESGRLAGAGVWPPAPAKDHPRTMPYNGMTPHI 334
233 KKHSHVYNTARGKIVNRPDALVEALASEHLQGYAGDVWYPPAPADHPRTMPRNMATVHY 292
335 SGTTLTAQARYAAGTRETILECFEGRPIRDEYLYVOGALAGTGAHSY 383
293 SGMTELEOKRIEDGVKDLERFFNHPEPQDDIIVASGRIA 333

RESULT 6
US-10-156-761-15009
; Sequence 15009, Application US/10156761
; Publication No. US20030119018A1

	Query Match	15.9%	Score 340.5;	DB 15;	Length 325;	
	Best Local Similarity	31.6%;	Pred. No. 8.2e-27;			
	Matches	94;	Conservative	46;	Mismatches 146;	Indels 11; Gaps 4
OY		70	TLVYTSKXDGSDSYFERELVDADVIYSQPFPAVLTPERIKAKKLKLTALTGISDHYD	129		
			: : : : : : : : : : : : : : : : : : :			
Dd		35	TVAATIDED-----PALAREAHYIIT-GIGP--VTAAHIAAPQLQCSHFDPDYD	85		
OY		130	LQSAIDRNVYAEV--TYCNSISVAEHVMMILSVRNYPSPHEWARGGWNIIADCVSHA	187		
			: : : : : : : : : : : : : : : : : : :			
Dd		86	LDARARGLLPCNINSGSGAEQOONAAEQFFALMLALAKOLVPNHTALVDADMPLRLQRSI	145		
OY		188	YDLFAMNHGTAAACRIGLAFLVRIPAFPVNHLHYDRHRLPSSEVEKELNIMTATREDMP	247		
			: : : : : : : : : : : : : : : : : : :			
Dd		146	TELSGRILGIYGLGHIGEVARRAVAFPMRITYAGREKYAGERERLGCARRHVIGDELIR	205		
OY		248	VCDVVYTNCLPHPETEMHINDETKLFERRAYIVNTAEGKLCDDRVAVALAESGRLAGYA	307		
			: : : : : : : : : : : : : : : : : : :			
Dd		206	TADVTTLHAPLTEATRHLLDADRLLALEPTFAFIWTAGALCIDODALLADELKAGALAG	265		
OY		308	GDVAFPPAPARDHMRTMYPNGMTPHSIGSTLTQAARAAGTRRELLECFEESBEPRIKD	364		
			: : : : : : : : : : : : : : : : : : :			
Dd		266	IDVEDPEBPPTSALKLLRAPNVVYLSSHAGVTRRETIYLRITALAAVQAADPAVGAEETPD	322		

```

:
:
: NUMBER OF SEQ ID NOS: 7059
:
: SOFTWARE: PatentIn ver. 3.0
:
: SEQ ID NO 4915
:
: LENGTH: 530
:
: type: prt
:
: ORGANISM: Corynebacterium glutamicum
:
: US-09-738-626-4915

```

```

Query Match      14.5%; Score 310.5; DB 10; Length 530;
Best Local Similarity 28.9%; Pred. No.2,3e-23;
Matches 103; Conservative 51; Mismatches 14; Indels 53; Gaps 12

QY      64  LESNGHTLVNYSDK-----DGPDSVFERELVD---ADVVISOPFWP 101
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      1  MSQGRPVLVLLADKLQSTVDALGDPAVEVRWDCGN---REPLDVAKEADALLVRS--A 55

QY      102 AYLLPERLAKKKNLKLTAIGSDSHVDLQSAIDRNVVAEYTVNSISVAEHVVMILS 161
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      56  TTVVAEYVLAAPNKKIYGRAGVGDNDNDIPATEGVVAVNAPSNIHSAEHAISLLS 115

QY      162 LVRRVYLPSEHARKGNVIADCVSHAYVLEMHVATVAGRIGLVLRRLRPFDVHLNYT 221
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      116 TARQIPAADALRLCEGEMKRS--FNGVYIEFGKTVIGVFGHIGQLFAORLAEFTTVAY 173

QY      222 DRHRLPESVEKELNLTWHTATREDMYPVCDVYVTLNCPRLPETEHMINDETLKFKGAYIV 281
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      174  DPVYNPAPA--AQLNVEL--VELDELMKSRSDPVTIHLPKETKTAGMDDQLAKSKSGQILLI 231

QY      282 NTAGKGLCDRRAVVALESGLRAGYAGVYFPPQAPKHPKHPMTMYNQMTHISGTTTLTA 341
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      232  NAAAGGLVDEQALDAIDISGIRAGGEFVYSTEPC--TSPSPFKRLQVVVYVTHLGLSTEEA 290

QY      342 QARRAAGVRELLIECFEGGRPIREDEYIYOGALAGTGAHSYSKGNATGS--SEENA 395
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      291  QDR--AGI-----DVADSVLKLAAEFVADAV-----NVSGRGGEVVA 327

```

```

RESULT 8
US-10-195-548-2
; Sequence 2, Application US/10195548
; Publication No. US20030008358A1
; GENERAL INFORMATION:
; APPLICANT: MIKIKO SUGA, Masakazu SUGIMOTO, Tsuyoshi OSUMI, Tsuyoshi
; APPLICANT: NAKAMATSU, Mataru HIBINO, Mika ITO
; TITLE OF INVENTION: METHOD OF PRODUCING L-SERINE BY FERMENTATION
; FILE REFERENCE: 02812
; CURRENT APPLICATION NUMBER: US/10/195,548
; PRIOR FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US/09/222,786
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: JP 10-3751
; PRIOR FILING DATE: 1998-01-12
; PRIOR APPLICATION NUMBER: JP 10-353513
; PRIOR FILING DATE: 1998-12-11
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 345
; TYPE: PRF
; ORGANISM: Corynebacterium glutamicum
US-10-195-548-2

Query Match          14.5%; Score 309.5; DB 15; Length 345;
Best Local Similarity 26.9%; Pred. No. 1.5e+23;
Matches 103; Conservative 51; Mismatches 149; Indels 53; Gaps 12

QY      .64 LESNGHTLVYTSDK-----DGPDSVFERELVD---ADVVISQPFWP 101
       : || ::|||               |||:         ||:::
Db      1 VSQGRPVVLADKLQSTVDALGDAVEVRWVGPN---RELLDAYKEADALLVRS--A 55

OY      102 AYLPELAKAKNKLALTAGISGDHVDLSAIDRNTVAEVTYCNSISVAEHVMMILS 161
       : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      56 TTYDAEVIAAPNKTIYGRAGVGLDNVDIPATAGVNAAPTSMHSACEHAISLLS 115

```

```

QY 16 LVNVLPRSHMRKGMGNINADCVSHAYZLEAMHNVTAGNIGLAVLRRLAPGEVHLHT 221
Db 116 TARQIPRAPAATITREGMTKRS--ENGVELFGKTVIGFEGHIGOLPRLAAEPFTTVAY 173
QY 222 DRRRLPESEVEKELNTWHTATREDMYPVCDVYVTLNCPRLPETHENHINDETUKLEFRGVYIV 281
Db 174 DRYANPARA-AQLNVEL-VELDELSRSDPVTIHLPKRTAGMFDQOLLAKSKGQII 231
QY 282 NTRARGLCRDVAVARLRESGLAGTAGDVWEPQAPKPHKPTKPTMYQGMTPRHISGTTLTA 341
Db 232 NAARGGLVDEQALAAIESGHAIRGAFPVYSTPEP-TDSPLEKPLQVVAVVPHILCASTEEA 290
QY 342 QARVATGTRILECFEFGPRIDEFLVYQGGALAGTGAHYSKGNATG--SEEA 395
Db 291 QDR-AGT-----DVADSVLKALAGEFADAV-----INSGRGAGEEVA 327

```

RESULT 9
US-10-195-548-12

Sequence 12, Application US/10195548
Publication No. US2003000838A1
GENERAL INFORMATION:
APPLICANT: MIKIKO SUGA, Masakazu SUGIMOTO, Tsuyoshi OSUMI, Tsuyoshih
APPLICANT: NAKAATSU, Wataru HIBINO, Mka ITO
TITLE OF INVENTION: METHOD OF PRODUCING L-SERINE BY FERMENTATION
FILE REFERENCE: JP812
CURRENT APPLICATION NUMBER: US/10/1955, 548
CURRENT FILING DATE: 2002-07-16
PRIOR APPLICATION NUMBER: US/09/322, 786
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: JP 10-3751
PRIOR FILING DATE: 1998-01-12
PRIOR APPLICATION NUMBER: JP 10-53513
PRIOR FILING DATE: 1998-12-11
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 530
TYPE: PRF
ORGANISM: *Brevibacterium flavum*
US-10-195-548-12

Query Match	14.5%;	Score 309.5;	DB 15;	Length 530;
Best Local Similarity	28.9%;	Pred. No. 3e-23;		
Matches 103; Conservative	51;	Mismatches 149;	Indels 53;	Gaps 12;

OY		64	LESNGHTLVNYSDK-----DGDSDVRELEVND---ADVISOPEMP	101
Dd		1	VSOQGRPVLLADKLAAOSTVDALGDAVEVRKWDDGN--RPELLDIYKRADALLVNS--A	55
OY		102	AYLTPERTIAKKNNLKLTALTAGISGDHVDLOSAIDRNVTVAEYTCNSISVAEHVMILS	161
Dd		56	TTVVAEYVTAAPNPNTKIYORAGLDGVNDVIPATEGAVMANPNTSIHSACEHAISLTL	115
OY		162	LVRNYLPSEHEARKGGMNIVADOVSXAUYDEAMHVAGTVAAGRGLAVIRLRLAPDYHLHT	221
Dd		116	TARQIPADAFLBREEMKRSS--FNGVELFTGTVLIGVGHHIGOLFQAOKLAEPETTIVAY	173
OY		222	DRHRLPESVEKELNLTWHTATREDMTPYCDDVYTLNCPLEHTEHMINDETLKFERRGAYTV	281
Dd		174	DPVANPNPARR-QOLANVEL-VELDELMSRSDFVTIHPKTRETMGMFDAQLAAKSKOIITI	231
OY		282	NFARGJLCDDRPAVARALESGRLAGYAAGVPOFAPKDHPRTMPPNMGMTPHISGTTLA	341
Dd		232	NAANGGLVDEBALDAIESGHTRGAGFPDYSTEPC-TDSPLTKLPVVYVTHPIHGASTEEEA	290
OY		342	QARYAAGTREILLECFEBEGRPTRIDEYLYIVOAGLACTGAHSYSGKNATGC--SEEAA	395
Dd		291	QDR--AGT-----DVADSVTLKLAGFEVDAY-----NVSGRGVBEVA	327

US-10-195-548-14
Sequence 14, Application US/10195548
Publication No. US20030008358A1
GENERAL INFORMATION:
APPLICANT: MIKIKO SUGA, Masakazu SUGIMOTO, Tsuyoshi OSUMI, Tsuyoshi
APPLICANT: NAKAMATSU, Wataru HIBINO, Mika ITO
TITLE OF INVENTION: METHOD OF PRODUCING L-SERINE BY FERMENTATION
FILE REFERENCE: JP812
CURRENT APPLICATION NUMBER: US/10/195,548
CURRENT FILING DATE: 2002-07-16
PRIOR APPLICATION NUMBER: US/09/222,786
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: JP 10-3751
PRIOR FILING DATE: 1998-01-12
PRIOR APPLICATION NUMBER: JP 10-353513
PRIOR FILING DATE: 1998-12-11
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 530
TYPE: PRF
ORGANISM: *Brevibacterium flavum*
US-10-195-548-14

Query Match	14.4%;	Score 307;	DB 15;	Length 530;
Best Local Similarity	30.5%;	Pred. No. 5.4e-23;		
Matches 94;	Conservative 49;	Mismatches 131;	Indels 34;	Gaps 9;

[illegible]

RESULT 10

RESULT 11
US-10-156-761-10267
: Sequence 10267, Application us/10156761
: Publication No. US20030119018A1
GENERAL INFORMATION:
: APPLICANT: OKURA, SATOSHI
: APPLICANT: IKEDA, HARDO
: APPLICANT: ISHIKAWA, JUN
: APPLICANT: HORIKAWA, HIROSHI
: APPLICANT: SHIBA, TADAYOSHI
: APPLICANT: SAKAKI, YOSHIYUKI
: APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDE
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: us/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02

76 DKDGPDSVFERELY-----DADVVISQPEWPAIYLTPERIA---KAKNLKALTAGIGSDH 127 QY

Db 51 BOEGEENALPEALLQGTDDIDILITQ-----PAPVTAPEFKIPKIKYGVLRGVEEN 104
Qy 128 VDLQSAIDRNTVAEVTYCNISVAEHVVMILSLVRNYLPSHEMARRGNINADCVSHA 187
Db 105 VNLQVANAARGVEVMTPCRNARSVAEFTVGMILAEMRINARSHDLRDKYR-KOSP HQ 163
Qy 188 Y--DLEMHVGTVAAGRIGLAVLRRLAFEDVHLHTDR-----HRLPEVEKELNLTWAT 241
Db 164 AIPELGKVVGLVGLGHTAOLVAGFLSGFTETIIFYDKYVAGHERYKVD-----S 214
Qy 242 REDVAPYCDVYTLNCLPHETEHMINDETLLKLFKRGAYIVNTARGKLCDDRAVARALEBSG 301
Db 215 LDELVRADVYLSHRALFPEENLINAHFALMKRSALITVATARGSLINEREMIDALRSG 274
Qy 302 RLAGYAGDVWFPQAPKDPWRTMPCYNGMTPHISGTTLLTA 341
Db 275 QIMGAAALDTFDEPLPDDSAFYTLNVTITPHIAGSTIDA 314

RESULT 14

US-10-177-293-92
; Sequence 92, Application US/10177293
; Publication No. US20030124128A1

GENERAL INFORMATION:

APPLICANT: Lillie, James
APPLICANT: Glatc, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Gannavarpu, Manjula
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Mertens, Maureen
APPLICANT: Myer, Vic
APPLICANT: Wang, Youzhen
APPLICANT: Xu, Yongyao
APPLICANT: Hoersch, Sebastian
APPLICANT: Monahan, John
APPLICANT: Meyers, Rachel E.
APPLICANT: Bast Jr., Robert C.
APPLICANT: Hortobagyl, Gabriel N.
APPLICANT: Pusztai, Lajos
APPLICANT: Meric, Aysegul
APPLICANT: Sahlu, Aysegul
TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-038
CURRENT APPLICATION NUMBER: US/10/177,293
CURRENT FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: US 60/299,887
PRIOR FILING DATE: 2001-06-21
PRIOR APPLICATION NUMBER: US 60/301,572
PRIOR FILING DATE: 2001-06-27
PRIOR APPLICATION NUMBER: US 60/306,501
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/325,002
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/362,585
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 60/xxx,xxx
PRIOR FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 506
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 92
LENGTH: 445
TYPE: PRT
ORGANISM: Homo sapiens
US-10-177-293-92

Query Match 12.9%; Score 275.5; DB 15; Length 445;
Best Local Similarity 28.0%; Pred. No. 7.9e-20;
Matches 90; Conservative 56; Mismatches 143; Indels 33; Gaps 7;

Qy 59 GLRPLAESNG--HTLVVTSKDGPDVSFE-----RELVDADYVI 95

Db 20 GIRPOI-MNGPLHPRPLVALLDGRCTYEMPIKDLATVAFCDASQTOEIHKEVLENVAG 78
Qy 96 SQPEWPAVLTPEERIAKAKNNLALATAGISDPVDLQSAIDRNTVAEVTYCNISVAEHV 155
Db 79 AMMYHTITLTREDEKFAFALRYIVAGSGYDWDVICKAAGELGIAVCNIPSAVEBTADST 138
Qy 156 VMILSLVRNLPFSHEMARRKG-----WNIDCVSHAYDLEMHVGTVAAGRIGLAVLR 210
Db 139 ICHILNLTNRMTWLQALREGTRVQSVQIREVNSGAAIRIRGETLGLGFRTGOAVAVR 198
Qy 211 LAPFDVHLHYTDHRHLPSVEKELNLTWHTATREDYPCDVYTLNCLPHETEHMINDET 270
Db 199 AKAGEFSYIFDPY-LDQIERSLGVQVYTLQDLITLQSDCVSLHCNLENNHHLINDPT 257
Qy 271 LKLFKRGAYIVNTARGKLCDDRAVARALESGRLAGYAGDVWFPQAPKDPWRTMPCYNG 329
Db 258 IKOMQGAFLVNAARGLVDEKALALQALKEGRIRGALDVEHSEPFSPAGPLKDAPIILI 317
Qy 330 MTPHIS--GTTLTQAQRTAAGT 349
Db 318 CTPTHAWYSEQASLEMRERAAAT 339

RESULT 15

US-10-205-823-76
; Sequence 76, Application US/10205823
; Publication No. US20030108963A1

GENERAL INFORMATION:

APPLICANT: Schlegel, Robert
APPLICANT: Monahan, John E.
APPLICANT: Endege, Wilson O.
APPLICANT: Gannavarpu, Manjula
APPLICANT: Gorbacheva, Bella
APPLICANT: Hoersch, Sebastian
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Womsey, Angela M.
APPLICANT: Glatc, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Anderson, Dustin
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF PROSTATE CANCER
FILE REFERENCE: MRI-044
CURRENT APPLICATION NUMBER: US/10/205,823
CURRENT FILING DATE: 2002-07-25
PRIOR APPLICATION NUMBER: 60/307,982
PRIOR FILING DATE: 2001-07-25
PRIOR APPLICATION NUMBER: 60/314,356
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/325,020
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: 60/341,746
PRIOR FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: 60/362,158
PRIOR FILING DATE: 2002-03-05
NUMBER OF SEQ ID NOS: 455
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 76
LENGTH: 985
TYPE: PRT
ORGANISM: Homo sapiens
US-10-205-823-76

Query Match 12.4%; Score 264; DB 15; Length 985;
Best Local Similarity 26.7%; Pred. No. 4.3e-18;
Matches 95; Conservative 61; Mismatches 160; Indels 40; Gaps 9;

Qy 32 PGGQLPMP-----KAIDFTPQQLIGS--VSGELGLRPLYESNG--HTLVVTSKDGPD 82
Db 526 PASQSLHTPHSPYEXVARRTGAPITVSTMLAPEFSIRPOI-MNGPLHPRPLVALLDGRDC 584
Qy 83 VFE-----RELVDADYVISQPPWPAVLTPEERIAKAKNNLALATLA 121

```

Db      585  TVEMPILKDLATVAFCDAOSTOEIHEKVLNEAVGAMMYHTITENREDLEKFKALRYIVRI 644
QY      122  GIGSDHYDLOSADIDRNTVAEVTYCNSISVAEHVMMILSLVRNYLPSHEWARKG--- 177
Db      645  GSGYDNYNITKAGELGIAVCNIPSAVEETADSTICHILNLYRNTWLYQALRGSTRVOS 704
QY      178  -WNIDCVSHAYDLEAMHVGTVAAGRIGLAVLRRLAPFDVHLHYTDHRRLPESYEKELNL 236
Db      705  VEOIRREYVAGSAGARIGETLIGIFGRTOQAVAVAKAFGSVIFYPDY-LQDGIERSLGV 763
QY      237  TWHATREDMYPCVDVVTLNCPLHPETEHMINDETLKFKGAYIVNTARGKLCDRDAVAR 296
Db      764  QRYTTLODDLXQSDCVSLHCNLNHNHLLINDEFTIKOMROGAFLVNAARGGLVDEKALAQ 823
QY      297  ALBSGRIAGYAGDVWFPQP-APKDPWRTPMRYNGMTPHIS--GTTLTAQARYAAGT 349
Db      824  ALKEGRIRGALADVHESEPFSAOGPLKDAPNLICTPHTAWYSEQASLEMRANAAT 879

```

Search completed: July 24, 2003, 20:44:25
 Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:43:38 ; Search time 99 Seconds
(without alignments)
1045.243 Million cell updates/sec

Title: US-09-996-008b-2

Perfect score: 2135
Sequence: 1 MAKVLCVLDPDVGYPKTY.....YSKGNATGSGSEAKPKKAV 401

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp_rodent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp_unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2127	99.6	401	2	Q93GV1 mycobacteri
2	2122	99.4	400	2	Q95V0 mycobacteri
3	1869	87.5	402	2	Q08375 moraxella s
4	1806	84.6	400	2	Q93GW3 paracoccus
5	1710	80.1	399	2	Q93JW1 hyphomicrob
6	1704	79.8	401	16	Q930E7 thizobium m
7	1511	70.8	398	2	Q9F7P9 uncultured
8	911	42.7	417	3	Q9Y790 mycosphaere
9	821	38.5	364	3	Q93968 candida boi
10	818	38.3	364	3	Q13437 candida boi
11	812	38.0	364	3	Q00498 candida met
12	803	37.6	376	3	Q08911 saccharomyc
13	781	36.6	374	16	Q8NYN1 staphylococ
14	777	36.4	374	16	Q99X44 staphylococ
15	520	24.4	248	10	Q8VX85 plinus pinas
16	502	23.5	199	10	Q8W520 zea mays (m

17	436	20.4	236	3	Q08988 saccharomyc
18	405.5	19.0	522	17	Q8TYK0 methanopyru
19	373	17.5	326	17	Q9Y8X7 aeropyrium p
20	352	16.5	333	17	Q9UYH9 pyrococcus
21	350.5	16.4	306	17	Q8UJ35 pyrococcus
22	348.5	16.3	528	16	Q8EN61 oceanobacil
23	347.5	16.3	323	17	Q8ZXX8 pyrobaculum
24	347	16.3	145	3	Q08987 saccharomyc
25	346.5	16.2	333	17	Q58256 pyrococcus
26	346	16.2	333	17	Q8U3T5 pyrococcus
27	345.5	16.2	307	17	Q9Y0M8 pyrococcus
28	340	15.9	354	16	Q8UJ26 pyrococcus
29	337.5	15.8	335	17	Q9UYR1 agrobacteri
30	336.5	15.8	307	17	Q50095 pyrococcus
31	332	15.6	323	17	Q28495 archaeglob
32	332	15.6	328	16	Q92YX6 thizobium m
33	326	15.3	376	17	Q58320 pyrococcus
34	323	15.1	449	13	Q9DEG7 brachydanio
35	322	15.1	540	17	Q8PM48 methanosc
36	319	14.9	322	2	Q59463 hyphomicrob
37	316	14.8	316	16	Q9CN80 pasteurella
38	313.5	14.7	534	16	Q8NM42 staphylococ
39	312.5	14.6	540	16	Q9KC99 bacillus ha
40	312	14.6	309	17	Q9HK29 thermoplasm
41	311.5	14.6	534	16	Q99TE0 staphylococ
42	310.5	14.5	530	16	Q8NOY7 corynebacte
43	309.5	14.5	530	16	Q8FPY9 corynebacte
44	308.5	14.4	334	16	Q92T15 thizobium m
45	308	14.4	360	16	Q8YEC6 bruceella me

ALIGNMENTS

RESULT 1
ID Q93GV1 PRELIMINARY: PRT: 401 AA.
AC Q93GV1
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Formate dehydrogenase (EC 1.2.1.2).
GN FDH.
GN Mycobacterium vaccae.
OC Bacteria; Actinobacteria; Actinobacteriales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1810;
RN [1]
RP SEQUENCE FROM N.A.
RA Galkin A., Kulakova L., Tishkov V., Esaki N., Soda K.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96158475; PubMed=8597552;
RA Galkin A., Kulakova L., Tishkov V., Esaki N., Soda K.;
RT "Cloning of formate dehydrogenase gene from a methanol-utilizing
bacterium Mycobacterium vaccae N10.";
RL Appl. Microbiol. Biotechnol. 44:479-483(1995).
DR EMBL: AB072394; BAB69476.1;
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH_C.
DR Pfam: PF00389; 2-Hacid_DH.1.
DR Pfam: PF02826; 2-Hacid_DH.C.1.
DR PROSITE: PS00065; D_2_HYDROXYACID_DH.1;
DR PROSITE: PS00670; D_2_HYDROXYACID_DH.2;
DR PROSITE: PS00671; D_2_HYDROXYACID_DH.3;
KW Oxidoreductase.
SQ SEQUENCE 401 AA: 44149 MW: 864821B4790AAD26 CRC64;
Query Match 99.6%; Score 2127; DB 2; Length 401;
Best Local Similarity 99.8%; Pred. No. 1.1e-16;
Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MAKVLCVLYDDPVDGCKRTYARDDLPRKIDHTPGGQILPTPKAIDFTPGQLGSSVSGELG 60
D 1 MAKVLCVLYDDPVDGCKRTYARDDLPRKIDHTPGGQILPTPKAIDFTPGQLGSSVSGELG 60
QY 61 RPYLESNGHTLVVTSDDGPDSPVERELVDADVVISOPFMPAYLTPERIAKAKNLKLTALT 120
D 61 RPYLESNGHTLVVTSDDGPDSPVERELVDADVVISOPFMPAYLTPERIAKAKNLKLTALT 120
QY 121 AGIGSDHVDLQSAIDRNVTAEVYTCNSISVAEHVMMILSLVNNYLPESHMARKGWN 180
D 121 AGIGSDHVDLQSAIDRNVTAEVYTCNSISVAEHVMMILSLVNNYLPESHMARKGWN 180
QY 181 ADCVSHAYDLEAMHVGTAAAGRIGLAVLRRLAPFDVHLHYTDRHRLPESVEKELNTLWHA 240
D 181 ADCVSHAYDLEAMHVGTAAAGRIGLAVLRRLAPFDVHLHYTDRHRLPESVEKELNTLWHA 240
QY 241 TREDMYPCDVVTLNCPHLPETEMINDETLKLFKRGAYIVNTARGKLCDDRAVARALES 300
D 241 TREDMYPCDVVTLNCPHLPETEMINDETLKLFKRGAYIVNTARGKLCDDRAVARALES 300
QY 301 GRLAGYAGDVWFPQAPKDPHWRTPYNGMTPHISGTTLLTAQARYAAGTREILECFEGR 360
D 301 GRLAGYAGDVWFPQAPKDPHWRTPYNGMTPHISGTTLLTAQARYAAGTREILECFEGR 360
QY 361 PIRDEYLIYOGGALAGTAHSHYSKGNATGSGSEAAKFKKAV 401
D 361 PIRDEYLIYOGGALAGTAHSHYSKGNATGSGSEAAKFKKAV 401

```

RESULT 2

Q9RSV0 PRELIMINARY; PRT; 400 AA.

```

AC Q9RSV0;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 23, Last annotation update)
DE NAD(+)-dependent formate dehydrogenase (EC 1.2.1.2) (Fragment).
OS Mycobacterium vaccae.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_Taxid=1810;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96158475; PubMed=8597552;
RA Gaikin A., Kulakova L., Tishkov V., Esaki N., Soda K.;
RT "Cloning of formate dehydrogenase gene from a methanol-utilizing
RT bacterium Mycobacterium vaccae N10."
RL Appl. Microbiol. Biotechnol. 44:479-483(1995).
DR HSSP: P33160; 2NAD.
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR Pfam: PF00389; 2-Hacid_DH.1.
DR Pfam: PF02826; 2-Hacid_DH.C.1.
DR PROSITE: PS00065; D_2_HYDROXYACID_DH.1; 1.
DR PROSITE: PS00670; D_2_HYDROXYACID_DH.2; 1.
DR PROSITE: PS00671; D_2_HYDROXYACID_DH.3; 1.
SQ SEQUENCE 400 AA; 44018 MW; 44228E88CA22BD9B CRC64;

```

Query Match 99.4%; Score 2122; DB 2; Length 400;
 Best Local Similarity 99.8%; Pred. No. 2.9e-166;
 Matches 399; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 2 AKVLCVLYDDPVDGCKRTYARDDLPRKIDHTPGGQILPTPKAIDFTPGQLGSSVSGELG 61
D 1 AKVLCVLYDDPVDGCKRTYARDDLPRKIDHTPGGQILPTPKAIDFTPGQLGSSVSGELG 60
QY 62 RPYLESNGHTLVVTSDDGPDSPVERELVDADVVISOPFMPAYLTPERIAKAKNLKLTALT 121
D 62 RPYLESNGHTLVVTSDDGPDSPVERELVDADVVISOPFMPAYLTPERIAKAKNLKLTALT 120
QY 121 AGIGSDHVDLQSAIDRNVTAEVYTCNSISVAEHVMMILSLVNNYLPESHMARKGWN 181
D 121 AGIGSDHVDLQSAIDRNVTAEVYTCNSISVAEHVMMILSLVNNYLPESHMARKGWN 180
QY 181 ADCVSHAYDLEAMHVGTAAAGRIGLAVLRRLAPFDVHLHYTDRHRLPESVEKELNTLWHA 240
D 181 ADCVSHAYDLEAMHVGTAAAGRIGLAVLRRLAPFDVHLHYTDRHRLPESVEKELNTLWHA 240
QY 241 TREDMYPCDVVTLNCPHLPETEMINDETLKLFKRGAYIVNTARGKLCDDRAVARALES 300
D 241 TREDMYPCDVVTLNCPHLPETEMINDETLKLFKRGAYIVNTARGKLCDDRAVARALES 300
QY 301 GRLAGYAGDVWFPQAPKDPHWRTPYNGMTPHISGTTLLTAQARYAAGTREILECFEGR 360
D 301 GRLAGYAGDVWFPQAPKDPHWRTPYNGMTPHISGTTLLTAQARYAAGTREILECFEGR 360

```

```

QY 182 DCVSHAYDLEAMHVGTAAAGRIGLAVLRRLAPFDVHLHYTDRHRLPESVEKELNTLWHA 241
D 182 DCVSHAYDLEAMHVGTAAAGRIGLAVLRRLAPFDVHLHYTDRHRLPESVEKELNTLWHA 240
QY 242 REDMYPCDVVTLNCPHLPETEMINDETLKLFKRGAYIVNTARGKLCDDRAVARALES 301
D 242 REDMYPCDVVTLNCPHLPETEMINDETLKLFKRGAYIVNTARGKLCDDRAVARALES 300
QY 241 REDMYPCDVVTLNCPHLPETEMINDETLKLFKRGAYIVNTARGKLCDDRAVARALES 300
D 241 REDMYPCDVVTLNCPHLPETEMINDETLKLFKRGAYIVNTARGKLCDDRAVARALES 300
QY 302 RLAGYAGDVWFPQAPKDPHWRTPYNGMTPHISGTTLLTAQARYAAGTREILECFEGR 361
D 302 RLAGYAGDVWFPQAPKDPHWRTPYNGMTPHISGTTLLTAQARYAAGTREILECFEGR 360
QY 361 IRDEYLIYOGGALAGTAHSHYSKGNATGSGSEAAKFKKAV 400
D 361 IRDEYLIYOGGALAGTAHSHYSKGNATGSGSEAAKFKKAV 400

```

RESULT 3

008375 PRELIMINARY; PRT; 402 AA.

```

AC 008375;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE NAD-dependent formate dehydrogenase (EC 1.2.1.2).
GN FDH.
OS Moraxella sp.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Moraxella.
OX NCBI_Taxid=479;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-1;
RA Gaikin A.G.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y13245; CAA73696.1; -.
DR HSSP: P33160; 2NAD.
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR Pfam: PF00389; 2-Hacid_DH.1.
DR Pfam: PF02826; 2-Hacid_DH.C.1.
DR PROSITE: PS00065; D_2_HYDROXYACID_DH.1; 1.
DR PROSITE: PS00670; D_2_HYDROXYACID_DH.2; 1.
DR PROSITE: PS00671; D_2_HYDROXYACID_DH.3; 1.
KW NAD; Oxidoreductase.
SQ SEQUENCE 402 AA; 44143 MW; C51CE1CE96BBB CRC64;

```

Query Match 87.5%; Score 1869; DB 2; Length 402;
 Best Local Similarity 85.7%; Pred. No. 1.9e-145;
 Matches 342; Conservative 26; Mismatches 31; Indels 0; Gaps 0;

```

QY 1 MAKVLCVLYDDPVDGCKRTYARDDLPRKIDHTPGGQILPTPKAIDFTPGQLGSSVSGELG 60
D 1 MAKVLCVLYDDPVDGCKRTYARDDLPRKIDHTPGGQILPTPKAIDFTPGQLGSSVSGELG 60
QY 61 RPYLESNGHTLVVTSDDGPDSPVERELVDADVVISOPFMPAYLTPERIAKAKNLKLTALT 120
D 61 RPYLESNGHTLVVTSDDGPDSPVERELVDADVVISOPFMPAYLTPERIAKAKNLKLTALT 120
QY 121 AGIGSDHVDLQSAIDRNVTAEVYTCNSISVAEHVMMILSLVNNYLPESHMARKGWN 180
D 121 AGIGSDHVDLQSAIDRNVTAEVYTCNSISVAEHVMMILSLVNNYLPESHMARKGWN 180
QY 181 ADCVSHAYDLEAMHVGTAAAGRIGLAVLRRLAPFDVHLHYTDRHRLPESVEKELNTLWHA 240
D 181 ADCVSHAYDLEAMHVGTAAAGRIGLAVLRRLAPFDVHLHYTDRHRLPESVEKELNTLWHA 240
QY 241 TREDMYPCDVVTLNCPHLPETEMINDETLKLFKRGAYIVNTARGKLCDDRAVARALES 300
D 241 TREDMYPCDVVTLNCPHLPETEMINDETLKLFKRGAYIVNTARGKLCDDRAVARALES 300
QY 301 GRLAGYAGDVWFPQAPKDPHWRTPYNGMTPHISGTTLLTAQARYAAGTREILECFEGR 360
D 301 GRLAGYAGDVWFPQAPKDPHWRTPYNGMTPHISGTTLLTAQARYAAGTREILECFEGR 360

```


Db 301 GRLAGYAGDVWFPQAPRDPHWRTPMNGMTPHISGTSLSQOTRKAAGTRELLECYFEGR 360
Qy 361 PIRDEYLLIVOGGALAGTGAHYSKGNATGSGSEAAKFKR 399
Db 361 PIRDEYLLIVOGGALAGTGAHYSKGNATGSGSEAAKFKR 399

RESULT 4

Q93GW3 PRELIMINARY; PRT; 400 AA.
ID 093GW3
AC 093GW3
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE NAD-dependent formate dehydrogenase.
GN Fdh.
OS Paracoccus sp. 12-A.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Paracoccus.
OX NCBI_TaxID=171192;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12-A;
RA Shinoda T., Satoh T., Mineki S., Iida M., Taguchi H.;
RT "Cloning, nucleotide sequence and expression in Escherichia coli of
RT the gene for formate dehydrogenase of Paracoccus sp. 12-A, a formate-
RT assimilating bacterium."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB071373; BAB64941.1;
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR Pfam: PF00389; 2-Hacid_DH; 1.
DR Pfam: PF02826; 2-Hacid_DH.C; 1.
DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.
DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
SQ SEQUENCE 400 AA; 43757 MW; 13195AED9AC299D8 CRC64;

Query Match 84.6%; Score 1806; DB 2; Length 400;
Best Local Similarity 83.2%; Pred. No. 3e-140;
Matches 333; Conservative 28; Mismatches 39; Indels 0; Gaps 0;

Qy 1 MAKVLCVLYDDPVNGYPRVTARDDLPKIDHPGGQILPTPKAIDFTPGQLGSGVSGELG 60
Db 1 MAKVLCVLYDDPVNGYPRVTARDDLPKIDHPGGQILPTPKAIDFTPGQLGSGVSGELG 60
Qy 61 RPYLESNGHTLVTSDDKGPDSVFERELVDADVVISOPFWPAYLTPERIAKAKNLKALT 120
Db 61 RNYLEAGCHELVTSDDKGPDSVFERELVDADVVISOPFWPAYLTPERIAKAKNLKALT 120
Qy 121 AGISGDHVDLQSAIDRNTVAEVTYCNISVAEHVMMILSLVRNYLPSHEMARKGNNI 180
Db 121 AGISGDHVDLQSAIDRNTVAEVTYCNISVAEHVMMILSLVRNYLPSHEMARKGNNI 180
Qy 181 ADCVSHAYDLEAMHVGTAAGRIGLAVLRRLAPFVNLHYTDNRHLPSEVEKELNLTHA 240
Db 181 ADCVSHAYDLEAMHVGTAAGRIGLAVLRRLAPFVNLHYTDNRHLPSEVEKELNLTHA 240
Qy 241 TREDMYPCADVNTLNCPLHPETEHIINDETLLKFRGAYIVNTARGKLCDDRAVAIALES 300
Db 241 SPKMFACDVNTLNCPLHPETEHIINDETLLKFRGAYIVNTARGKLCDDRAVAIALES 300
Qy 301 GRLAGYAGDVWFPQAPRDPHWRTPMNGMTPHISGTSLSQOTRKAAGTRELLECYFEGR 360
Db 301 GRLAGYAGDVWFPQAPRDPHWRTPMNGMTPHISGTSLSQOTRKAAGTRELLECYFEGR 360
Qy 361 PIRDEYLLIVOGGALAGTGAHYSKGNATGSGSEAAKFKR 400
Db 361 PIRDEYLLIVOGGALAGTGAHYSKGNATGSGSEAAKFKR 400

RESULT 5

Q93UW1 PRELIMINARY; PRT; 399 AA.
ID 093UW1
AC 093UW1
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE NAD+ dependent formate dehydrogenase.
GN Fdh.
OS Rhizobium sp. JCI7.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=142666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCI7;
RA Tanaka Y., Yoshida T., Watanabe K., Mitsuura T.;
RT "Characterization, gene cloning and expression of NAD+ dependent
RT formate dehydrogenase from a methylotrophic bacterium Rhizobium
RT sp. JCI7."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB051073; BAB54449.1;
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR Pfam: PF00389; 2-Hacid_DH; 1.
DR Pfam: PF02826; 2-Hacid_DH.C; 1.
DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
SQ SEQUENCE 399 AA; 44044 MW; 853BAD2C635864DC CRC64;

Query Match 80.1%; Score 1710; DB 2; Length 399;
Best Local Similarity 78.2%; Pred. No. 2.4e-132;
Matches 312; Conservative 42; Mismatches 45; Indels 0; Gaps 0;

Qy 1 MAKVLCVLYDDPVNGYPRVTARDDLPKIDHPGGQILPTPKAIDFTPGQLGSGVSGELG 60
Db 1 MAKVLCVLYDDPVNGYPRVTARDDLPKIDHPGGQILPTPKAIDFTPGQLGSGVSGELG 60
Qy 61 RPYLESNGHTLVTSDDKGPDSVFERELVDADVVISOPFWPAYLTPERIAKAKNLKALT 120
Db 61 RPYLESNGHTLVTSDDKGPDSVFERELVDADVVISOPFWPAYLTPERIAKAKNLKALT 120
Qy 121 AGISGDHVDLQSAIDRNTVAEVTYCNISVAEHVMMILSLVRNYLPSHEMARKGNNI 180
Db 121 AGISGDHVDLQSAIDRNTVAEVTYCNISVAEHVMMILSLVRNYLPSHEMARKGNNI 180
Qy 181 ADCVSHAYDLEAMHVGTAAGRIGLAVLRRLAPFVNLHYTDNRHLPSEVEKELNLTHA 240
Db 181 ADCVSHAYDLEAMHVGTAAGRIGLAVLRRLAPFVNLHYTDNRHLPSEVEKELNLTHA 240
Qy 241 TREDMYPCADVNTLNCPLHPETEHIINDETLLKFRGAYIVNTARGKLCDDRAVAIALES 300
Db 241 SLESITKACDVNTLNCPLHPETEHIINDETLLKFRGAYIVNTARGKLCDDRAVAIALES 300
Qy 301 GRLAGYAGDVWFPQAPRDPHWRTPMNGMTPHISGTSLSQOTRKAAGTRELLECYFEGR 360
Db 301 GRLAGYAGDVWFPQAPRDPHWRTPMNGMTPHISGTSLSQOTRKAAGTRELLECYFEGR 360
Qy 361 PIRDEYLLIVOGGALAGTGAHYSKGNATGSGSEAAKFKR 399
Db 361 PIRDEYLLIVOGGALAGTGAHYSKGNATGSGSEAAKFKR 399

RESULT 6

Q93OE7 PRELIMINARY; PRT; 401 AA.
ID 093OE7
AC 093OE7
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Probable NAD-dependent formate dehydrogenase.
GN Rhizobium meliloti (Sinorhizobium meliloti).
OC Plasmid pSymA (megaplasmid 1).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

Query Match	79.8%	Score 1704	DB 16	Length 401
Best Local Similarity	78.4%	Pred. NO. 7.4e-132		
Matches 312	Conservative 39	Mismatches 47	Indels 0	Gaps 0

```

RESULT 7
Q9F7P9
ID Q9F7P9      PRELIMINARY:      PRT;      398 AA.
AC Q9F7P9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Predicted NAD-dependent formate dehydrogenase.
OS uncultured proteobacterium EBA331A08.
OC Bacteria; Proteobacteria; environmental samples.
CX NCBI_taxid=133804;

```

Query Match	70.8%	Score 1511	DB 2	Length 398
Best Local Similarity	69.0%	Pred. No. 5.6e-116		
Matches 274	Conservative 54	Mismatches 69	Indels 0	Gaps 0
RA	Jovanovich S.B., Gates C.M., Feldman R.A., Spudich J.L., Spudich E.N.,			
DR	Delong E.F.:			
RT	"Bacterial rhodopsin: evidence for a new type of phototrophy in the			
RL	sea."			
DR	Science 289:1902-1906(2000).			
DR	EMBL: AF279106; AAC10470.1; -			
DR	HSSP: P33160; 2NAD			
DR	InterPro: IPR006139; 2-Hacid_DH			
DR	InterPro: IPR006140; 2-Hacid_DH_C			
DR	Pfam: PF0389; 2-Hacid_DH; 1.			
DR	Pfam: PF02826; 2-Hacid_DH_C; 1.			
DR	PROSITE: PS00065; D_2_HYDROXYCID_DH_1; 1.			
DR	PROSITE: PS00671; D_2_HYDROXYCID_DH_3; 1.			
DR	SEQUENCE 398 AA; 43992 MW; 7C2EC76FD5723C4C CRC64;			
QY	3 KVLCTLYDDPYDGYKTYARDLPKIDHYPGQGLPFRKAIDFPFGQLGSVSGELGRP 62			
DB	2 KILCTLYDDPTKGMERYARDLPKIDYPPQGMPLPSKSIDFTFGELGCVSGELGRK 61			
QY	63 YLESNGHTLVYTSKDKCPDSVFPERLVADAVYISQPFPPAYLTPERIAKAKLALING 122			
DB	62 FLEDAHGHTLVYTSDDGDGCEADKTLVADYISQPFPPYLTTRKMTAPMLKMAITAG 121			
QY	123 IGSDDHVDQSADIRNVTAEVYTCISVAEHHVMILSYVNYLRPSHEMARKGMNTAD 182			
DB	122 IGSDDHVDLQAAAMDNSVDVETVYTCRSRSVAELIYVMILSMWVDITQRIHYEGGMNTAD 181			
QY	183 CVSHAYDLEAMHGVTAAGRIGLAVLRRLAPRDVHLHYTHDRHRLPESVEKELNTLWHAIR 242			
DB	182 AVORSYDVEGHNVGTVAAGRIGIDMLRKMKPFVDVHLHYFDLHKLSDEIEAELNTLYHDSV 241			
QY	243 EDMPYVCVNVTLNCLHPETEHEMIDETLKLKRGATVYNTARGKLCRDARALRESGR 302			
DB	242 ESLAVAVCVYVNIISCLHRTKTEHLPEDDEIKSKKKRATYIINTARGKICDKDAIRLESQ 301			
QY	303 LAGVAGDVWFPQAPKDPHWRTPYNGMTPHISGTTLLTAQARYAAGREILCEFFEGRI 362			
DB	302 LSGVAGDVWFPQAPNDVHWRTPMHHGMTPHSGTSLSAQRTYRAGVREILCEYFAGEPI 361			
QY	363 RDEVLLIYOGGALAGTGAHSYSGNATGSGSEAKRKK 399			
DB	362 RDPYLLIYONGDLAGKGAHSYTKGATPDSSEAKRKK 398			
RESULT 8				
ID	Q9Y790	PRELIMINARY:	PRT:	417 AA.
AC	Q9Y790:			
DT	01-NOV-1999 (TREMBLrel. 12, Created)			
DR	01-NOV-1999 (TREMBLrel. 12, Last sequence update)			
DR	01-OCT-2002 (TREMBLrel. 22, Last annotation update)			
DE	NAD-dependent formate dehydrogenase (EC 1.2.1.2).			
OS	Mycosphaerella graminicola (Septoria tritici).			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina;			
OC	Dottidomycetes et Chaetothyriomycetes Incertae sedis;			
OC	Mycosphaerellaceae; Mycosphaerella.			
OX	NCBI_TaxID=54734;			
OX	NCBI_TaxID=54734;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Stritt;			
RA	Keon J.P.R., Bailey A.M., Hargreaves J.A.;			
RT	"cDNA clone expressed when utilizing ammonium ions as the major source			
RT	of nitrogen."			
RL	Submitted (Jan-1999) to the EMBL/Genbank/DBJ databases.			
DR	EMBL: AF131482; AAD3831.1; -			
DR	HSSP: P33160; 2NAD			
DR	InterPro: IPR006139; 2-Hacid_DH			
DR	InterPro: IPR006140; 2-Hacid_DH_C			
DR	Pfam: PF0389; 2-Hacid_DH; 1.			

DR Pfam; PF02826; 2-Hacid_DH.C; 1.
 DR PROSITE; PS0065; D_2_HYDROXYACID_DH.1; 1.
 DR PROSITE; PS00670; D_2_HYDROXYACID_DH.2; 1.
 DR PROSITE; PS00671; D_2_HYDROXYACID_DH.3; 1.
 KW Oxidoreductase.
 SQ SEQUENCE 417 AA; 45627 MW; 916E0BC35AF9EDD CRC64;

Query Match 42.7%; Score 911; DB 3; Length 417;
 Best Local Similarity 49.7%; Pred. No. 1.5e-66;
 Matches 196; Conservative 51; Mismatches 105; Indels 42; Gaps 7;

QY 3 KVLCLVDDPYDGYPKTYARDLPRIDHYPGQQLIPTRKAIIDFTPGQLGSGVSGELGRP 62
 DB 51 KVLVLVDG-----HEHAQOE-----PRLLGTENTLGRK 81
 QY 63 YLESGHRLVYTSKDGSDSFERELVDADVISOFPMPAYLTPERIAKAKLALTAG 122
 DB 82 WIEOGHRLVYTSKDEGNSKFDQELVDAEVIITPPFPGYLTARLAKAKKLKATVAG 141
 QY 123 IGSFHVLDQSAIDNR--VTVAEVYCNISVAEHVMMILSVRYLPSHEMARKGN 180
 DB 142 IGSFHVLDNANKNKGITVAEYVGSNVAEHWMTLVLRNFPVPAHQIAGAOWNV 201
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPDV-HIHYDRRLRLESYKEKELNTW 239
 DB 202 AAVAKNEYDEGKVGTVAVAGRIGERVLRLKPFCKELLYFDYQALPVEYKEIGCRV 261
 QY 240 ATREDMPVCDVYVTLNCPHPETEMINDETLKEFKGAYVNTARGKLCRDVAARALE 299
 DB 262 DTEELVLAQCVVIVINCPLHETKRLFKELISKKKKSWLVNTRGALIVKEEVAALK 321
 QY 300 SGRLAGYAGDVWFPAPARHPMPY-----NGMPTPISTGLTQAARYAGTREIL 353
 DB 322 FGQLRGIGGDVWFPKRPADHPERTASTYWGNGNAMPHMSGISIDQARIAGTKAIL 381
 QY 354 ECFEGR-PIRDEYLIYOGG 386
 DB 382 DSYSFGREDYRPEDLIYHKGDYA---TKAYGORN 412

RESULT 9

093968 PRELIMINARY; PRT; 364 AA.
 AC 093968;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Formate dehydrogenase (EC 1.2.1.2).
 GN FDI.
 OS Candida boidinii (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5477;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCYC 1513;
 RA Labrou E.;
 RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AJ011046; CAA09466.2; -
 DR HSSP; P33160; 2NAC.
 DR InterPro; IPR006139; 2-Hacid_DH.
 DR InterPro; IPR006140; 2-Hacid_DH.C.
 DR Pfam; PF00389; 2-Hacid_DH.C; 1.
 DR Pfam; PF02826; 2-Hacid_DH.C; 1.

DR PROSITE; PS0065; D_2_HYDROXYACID_DH.1; 1.
 DR PROSITE; PS00670; D_2_HYDROXYACID_DH.2; 1.
 DR PROSITE; PS00671; D_2_HYDROXYACID_DH.3; 1.
 KW Oxidoreductase.
 SQ SEQUENCE 364 AA; 40266 MW; 34B727B7493D7B3F CRC64;

Query Match 38.5%; Score 821; DB 3; Length 364;
 Best Local Similarity 45.1%; Pred. No. 3.1e-59;
 Matches 171; Conservative 57; Mismatches 113; Indels 38; Gaps 6;

QY 3 KVLCLVDDPYDGYPKTYARDLPRIDHYPGQQLIPTRKAIIDFTPGQLGSGVSGELGRP 62
 DB 2 KVLVLVDG-----AGKHADE-----EKLYGTENKLGIAN 32
 QY 63 YLESGHRLVYTSKDGSDSFERELVDADVISOFPMPAYLTPERIAKAKLALTAG 122
 DB 33 WIKDGHRLVYTSKDEGNSVLDQIIPADIIITPPFPGYLTARLAKAKKLKATVAG 92
 QY 123 IGSFHVLDQ--SAIDRNTVAEVYCNISVAEHVMMILSVRYLPSHEMARKGN 180
 DB 93 VGSFHVLDIYINQTKKISVLEVTGSNVAEHWMTLVLRNFPVPAHQIAGAOWNV 152
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPDV-VHLHYDRRLRLESYKEKELNTW 239
 DB 153 AAIKADVDIEGKTATIGAGRIIGYRVLRLVFPNPKELLYDYQALPKDAEKGARV 212
 QY 240 ATREDMPVCDVYVTLNCPHPETEMINDETLKEFKGAYVNTARGKLCRDVAARALE 299
 DB 213 ENIELVLAQDVIVYVNAPLHAGTKLINKELLSKFKKGMVNTARGALICVAEDVAAL 272
 QY 300 SGRLAGYAGDVWFPAPARHPMPY-----PYNMTPISTGLTQAARYAGTREIL 354
 DB 273 SGQLRGIGGDVWFPAPARHPMPY-----PYNMTPISTGLTQAARYAGTREIL 332
 QY 355 ECFEGR-PIRDEYLIYOGG 372
 DB 333 SFTGKFDYRPQDITLLNG 351

RESULT 10

013437 PRELIMINARY; PRT; 364 AA.
 AC 013437;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE NAD-dependent formate dehydrogenase (EC 1.2.1.2).
 GN FDI OR FDI3.
 OS Candida boidinii (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5477;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S2;
 RA Sakai T.; Muranoto A.P.; Konishi T.; Iwamatsu A.; Kato N.;
 RA MEDLINE-97369815; PubMed-9226256;
 RA "Regulation of the formate dehydrogenase gene, FDI1, in the
 RT methylotrophic yeast Candida boidinii and growth characteristics of an
 RT FDI1-disrupted strain on methanol, methylamine, and choline.";
 RL J. Bacteriol. 179:4480-4485(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC32195;
 RA Slusarczyk H.;
 RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF004096; AAC49766.1; -
 DR EMBL; AJ245934; CAB54834.1; -

DR HSSP; P33160; 2NAC.
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam: PF00389; 2-Hacid_DH; 1.
 DR Pfam: PF02826; 2-Hacid_DH.C; 1.
 DR PROSITE; PS00065; D_2_HYDROXYACID_DH_1; 1.
 DR PROSITE; PS00670; D_2_HYDROXYACID_DH_2; 1.
 DR PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.
 DR Oxidoreductase.
 SO SEQUENCE 364 AA; 40370 MW; 1B30982EDD5877EB CRC64;

Query Match 38.3%; Score 818; DB 3; Length 364;
 Best Local Similarity 44.9%; Pred. No. 5.5e-59;
 Matches 170; Conservative 57; Mismatches 114; Indels 38; Gaps 6;

QY 3 KVLCLYDDPYDGYPKTYARDLPRKIDHYPEGQLPFPKALDFTPGQLGSGVSGELGRP 62
 Db 2 KIVLVLD-----AGKHADE-----EKLYGCTENKLGIAN 32
 QY 63 YLENGHTLVTSKDDGDSFERELVDADVVISQFPMPAYLTPERLAKANKLALTAG 122
 Db 33 WLKQGHLLTTSDEGETSELDKHIPDADIIITPPHAYITKERLDKANKLKVYVAG 92
 QY 123 IGSDDVLDQ--SAIDRNTVAEVTYCNSSISAEHVMMILSLVNYLPSHEMARKGMI 180
 Db 93 VGSHTIDLDYINQTKKISVLETVGSNVSAEHVMTMLVNFVPAHQIINHMEV 152
 QY 181 ADCVSHAYDLEAMHGVAAAGTGLAVLRRLAPD-VHLHYTDNRHRLPESVEKELNTW 239
 Db 153 AAIAKDAYDIEGKTATIGAGTIGRYRLERLPPNPPELLLYDQALPKEAEKVGARV 212
 QY 240 ATREDMTPVCDVYVTLNCPLEPETHMINDETLKFKGAVYVNTARGKLCDDRAVARALE 299
 Db 213 ENIELVAQADIVYVNPALHAGTGKGLINKELLSFKKGAMVNTARGAICVAEDVAAALE 272
 QY 300 SGRLAGYAGDVWFPQAPKDPKHPMTM----PYNGMTPHISGTLTQAARVAACTREILE 354
 Db 273 SGQLRGGYGVWFPQAPKDPKHPMDMKKYGAGNAMPHTSGTLLDQTRAEKTKNILE 332
 QY 355 CFEEGR-PIRDEYLIYOGG 372
 Db 333 SFTGKFDYRPQDITILLG 351

RESULT 11
 ID 000498 PRELIMINARY; PRT; 364 AA.
 AC 000498;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
 DE NAD-dependent formate dehydrogenase.
 OS Candida methyllica.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=45353;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96009885; PubMed=7557425;
 RA Allen S.J., Holbrook J.J.;
 RT "Isolation, sequence and overexpression of the gene encoding NAD-
 dependent formate dehydrogenase from the methylotrophic yeast Candida
 methyllica.";
 RT Gene 162:99-104(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Hall L.;
 RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X81129; CAA57036.1; -
 DR HSSP; P33160; 2NAC.
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam; PF00389; 2-Hacid_DH; 1.

DR Pfam; PF02826; 2-Hacid_DH.C; 1.
 DR PROSITE; PS00065; D_2_HYDROXYACID_DH_1; 1.
 DR PROSITE; PS00670; D_2_HYDROXYACID_DH_2; 1.
 DR PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.
 DR NAD.
 SO SEQUENCE 364 AA; 40344 MW; D83DD1FA9D5EA303 CRC64;

Query Match 38.0%; Score 812; DB 3; Length 364;
 Best Local Similarity 44.6%; Pred. No. 1.7e-58;
 Matches 169; Conservative 57; Mismatches 115; Indels 38; Gaps 6;

QY 3 KVLCLYDDPYDGYPKTYARDLPRKIDHYPEGQLPFPKALDFTPGQLGSGVSGELGRP 62
 Db 2 KIVLVLD-----AGKHADE-----EKLYGCTENKLGIAN 32
 QY 63 YLENGHTLVTSKDDGDSFERELVDADVVISQFPMPAYLTPERLAKANKLALTAG 122
 Db 33 WLKQGHLLTTSDEGETSELDKHIPDADIIITPPHAYITKERLDKANKLKVYVAG 92
 QY 123 IGSDDVLDQ--SAIDRNTVAEVTYCNSSISAEHVMMILSLVNYLPSHEMARKGMI 180
 Db 93 VGSHTIDLDYINQTKKISVLETVGSNVSAEHVMTMLVNFVPAHQIINHMEV 152
 QY 181 ADCVSHAYDLEAMHGVAAAGTGLAVLRRLAPD-VHLHYTDNRHRLPESVEKELNTW 239
 Db 153 AAIAKDAYDIEGKTATIGAGTIGRYRLERLPPNPPELLLYDQALPKEAEKVGARV 212
 QY 240 ATREDMTPVCDVYVTLNCPLEPETHMINDETLKFKGAVYVNTARGKLCDDRAVARALE 299
 Db 213 ENIELVAQADIVYVNPALHAGTGKGLINKELLSFKKGAMVNTARGAICVAEDVAAALE 272
 QY 300 SGRLAGYAGDVWFPQAPKDPKHPMTM----PYNGMTPHISGTLTQAARVAACTREILE 354
 Db 273 SGQLRGGYGVWFPQAPKDPKHPMDMKKYGAGNAMPHTSGTLLDQTRAEKTKNILE 332
 QY 355 CFEEGR-PIRDEYLIYOGG 372
 Db 333 SFTGKFDYRPQDITILLG 351

RESULT 12
 ID 008911 PRELIMINARY; PRT; 376 AA.
 AC 008911;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
 DE Chromosome XY reading frame ORF YOR388C.
 GN FDI1 OR YOR388C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetales; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Delius H., Hebling U., Hofmann B.;
 RL submitted (Jul-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MIPS;
 RL Submitted (Jul-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z75296; CAA99720.1; -
 DR HSSP; P33160; 2NAD.
 DR SGD; S0005915; FDI1.
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam; PF00389; 2-Hacid_DH; 1.
 DR Pfam; PF02826; 2-Hacid_DH.C; 1.
 DR PROSITE; PS00670; D_2_HYDROXYACID_DH_2; 1.
 DR PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.
 SO SEQUENCE 376 AA; 41714 MW; 67ECD46F9DDC2A02 CRC64;

Query Match 37.6%; Score 803; DB 3; Length 376;

Best Local Similarity 44.7%; Pred. No. 9.9e-58;
Matches 181; Conservative 53; Mismatches 115; Indels 56; Gaps 9;

```

QY 3 KVLCLVLDPPVDPYKPKTYARDLKRIDHYPGQQLIPKAIIDFTPGQLSGVSGELGRP 62
DB 5 KVLVLVE-----GKRLHEBOE-----KILGCIENELGIRN 35
QY 63 YLSENGHTLVYTSKD-GPDSVFERELVDADVISOFPWPAVLPERIAKAKNLKALTA 121
DB 36 FIEQGYELVTTIDKDEPSTVDRELKDAEIVITTPPEPAVISNRILAEPANIKLCYTA 95
QY 122 GIGSDHDLQASIDRNVVAEYTCNSIVAEHVMYMLISLRNTLPSEHNAKRGWYIA 181
DB 96 GVGSDHDLQASIDRNVVAEYTCNSIVAEHVMYMLISLRNTLPSEHNAKRGWYIA 155
QY 182 DCVSHAYDLEAMHVGVAAGRIGLAVLRRLAPFD-VHLHYDRHRRLPESVERKELMTWHA 240
DB 156 GVAKNEYDLEKDIISTVGAGRIGYRLERLVAFNPKLLYYDYQELPAEALNRLN---EA 212
QY 241 TR-----EDMYPVGDVYVTLNCLPHEPEHMINDETLKFKRGAYIVNTR 285
DB 213 SKLNGRQDIYORVEKLEDMYAGSDVYVTLNCLPHEPEHMINDETLKFKRGAYIVNTR 272
QY 286 GKLCRDPAVARALSGRLAGYAGDVWFPQAPAKDHPMRTMPY-----NGMTPHISGTTLT 340
DB 273 GAICVAEDVAAYVSGKLAGYGGDVWDKQAPAKDHPMRTMKNKHVGNAMKTVHISGTSILD 332
QY 341 AQAARYAGTRILECFEFGGR-PIIDEYLIYOGGALAGTGAHSYSK 384
DB 333 AQRKYAGGVKNILNSYFSKFKFDYRQDIYONGSYA-TRAYGOKK 376

```

RESULT 13

```

Q8NVN1 PRELIMINARY; PRT; 374 AA.
ID Q8NVN1
AC Q8NVN1;
DT 01-OCT-2002 (TREMblrel. 22, Created)
DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE MAD-dependent formate dehydrogenase.
GN FDH OR MW0151.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
acquired MRSA.";
RL Lancet 358:1819-1827 (2002).
DR EMBL; AP004822; BAB94016.1;
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR Pfam: PF00389; 2-Hacid_DH.1.
DR Pfam: PF02826; 2-Hacid_DH.C.1.
DR PROSITE; PS00065; D_2_HYDROXYACID_DH.1; 1.
DR PROSITE; PS00670; D_2_HYDROXYACID_DH.2; 1.
KW Complete proteome.
SQ SEQUENCE 374 AA; 41945 MW; 39BE6D84FD58D587 CRC64;

```

Query Match 36.6%; Score 781; DB 16; Length 374;
Best Local Similarity 45.6%; Pred. No. 6.3e-56;
Matches 159; Conservative 64; Mismatches 104; Indels 22; Gaps 6;

```

QY 40 PKAIDFTPGQLSGVSGELGRPYLESNGHTLVYTSKDGPDSVFERELVDADVISOFP 99
DB 41 PEAVGQENQLNTRKA-IGLKTFLERGHETIAD-NGED--LDKHLPMDDVITISAPF 96
QY 100 WPAVLTPEERIAKAKNLKALITAGIGSDHVDLSAIDRNVVAEYTCNSIVAEHVMYMI 159
DB 157 LILRNVEGHRQSVGEWMNLQVGNHAEHLQHTKIGIFGFRIGOLVAERLAPPVTLQ 216

```

```

DB 97 YPAVYTRERIEKAPNLKALITAGVGSODHDLAAASEHNIIGVEYTSNTVSAEHAVMOL 156
QY 160 LSLVANYLPSEHNAKRGWYIAIDCVSHAYDLEAMHVGVAAGRIGLAVLRRLAPFDVHLH 219
DB 157 LILRNVEGHRQSVGEWMNLQVGNHAEHLQHTKIGIFGFRIGOLVAERLAPPVTLQ 216
QY 220 YTD-----RHRLPESVERKELMTWATREDMYPVGDVYVTLNCLPHEPEHMINDETLKLF 274
DB 217 HYDPIQDHRKSKFV-----SFDELSTSDALITAPLPEEDNLFQKDYLSRM 266
QY 275 KRGAIVYNTANGKLCRDPAVARALSGRLAGYAGDVWFPQAPAKDHPMRTMPYNGMTPHI 334
DB 267 KKHSLVYNTANGKIYNRRALVALEASHLQAGDVWFPQAPADHPMRTMKNKTVHISGTSILD 326
QY 335 SGTTLTAQARYAGTRILECFEFGGRPIIDEYLIYOGGALAGTGAHSYSK 383
DB 327 SGMTELAQRKIEDGVKILERFNEPQDKDIIVASGRIA---SKSYT 372

```

RESULT 14

```

Q99X44 PRELIMINARY; PRT; 374 AA.
ID Q99X44
AC Q99X44;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE MAD-dependent formate dehydrogenase.
GN FDH OR SAV0177 OR SAV0171.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohba T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cul L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanemori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-U Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Fuyuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus.";
RL Lancet 357:1225-1240 (2001).
DR EMBL; AP003358; BAB56339.1;
DR EMBL; AP003129; BAB41392.1;
DR HSSP; P33160; 2NAD.
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR Pfam: PF00389; 2-Hacid_DH.1.
DR Pfam: PF02826; 2-Hacid_DH.C.1.
DR PROSITE; PS00065; D_2_HYDROXYACID_DH.1; 1.
DR PROSITE; PS00670; D_2_HYDROXYACID_DH.2; 1.
KW Complete proteome.
SQ SEQUENCE 374 AA; 41931 MW; 4B31FC87DBF2864D CRC64;

```

Query Match 36.4%; Score 777; DB 16; Length 374;
Best Local Similarity 45.3%; Pred. No. 1.4e-55;
Matches 158; Conservative 65; Mismatches 104; Indels 22; Gaps 6;

```

QY 40 PKAIDFTPGQLSGVSGELGRPYLESNGHTLVYTSKDGPDSVFERELVDADVISOFP 99
DB 41 PEAVGQENQLNTRKA-IGLKTFLERGHETIAD-NGED--LDKHLPMDDVITISAPF 96
QY 100 WPAVLTPEERIAKAKNLKALITAGIGSDHVDLSAIDRNVVAEYTCNSIVAEHVMYMI 159
DB 97 YPAVYTRERIEKAPNLKALITAGVGSODHDLAAASEHNIIGVEYTSNTVSAEHAVMOL 156
QY 160 LSLVANYLPSEHNAKRGWYIAIDCVSHAYDLEAMHVGVAAGRIGLAVLRRLAPFDVHLH 219
DB 157 LILRNVEGHRQSVGEWMNLQVGNHAEHLQHTKIGIFGFRIGOLVAERLAPPVTLQ 216

```

```

QY 220 YTO-----NHRLEESVEKELNTLMHATREDMPQCDVATLINCPLPHEPENHINETKLF 274
Db 217 HYDPIQODHKLKFEV-----SFDVLSSSDAITIHAPLPEPDNLFDRDVLSSRM 266
QY 275 KRASYIVNARAGLCORDAVRALESGLRAGYAGGVWPPRARPDHRYTPYVGMPRI 334
Db 267 KRHSYLVNARKGIYVRDLVRLALSNSHQIAGAGVMPYQFAPADHRIKTIYRNATVHY 328
QY 335 SGTLLIQAQRYANAGTFLIECEPFBGRPIRDEYLIYOGALAGTGHSTYS 383
Db 327 SGTLLLEAKRIEDGVADILEREFFNHRPDCKDIIVASSRIA---SKST 372

```

RESULT 15

ID			PRELIMINARY;	PRT;	248 AA.
AC	O8YX85:				
D_T	01-MAR-2002 (TREMBLrel_20,		Created)		
D_T	01-MAR-2002 (TREMBlrel_20,		Last sequence update)		
D_T	01-MAR-2003 (TREMBlrel_23,		Last annotation update)		
DE	Putative NAD-dependent formate dehydrogenase (EC 1.2.1.2)		(Fragment).		
OS	Pinus pinaster (Maritime pine).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
CC	Spermatophytes; Coniferopsida; Coniferales; Pinaceae; Pinus.				
OX	NCBI_TaxId=71647;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	TISSUE=Root;				
RC	Dubos C.; Plomlon C.;				
RT	"Identification of water-deficit responsive genes in maritime pine"				
RL	(Pinus pinaster Ait.) roots using non radioactive cDNA-AFLP."				
DR	Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.				
DR	EMBL: AJ300721; CA683306.1;				
DR	InterPro: IPR006140; 2-Hacid_DH_C.				
DR	Pfam: PF02826; 2-Hacid_DH_C.1.				
DR	PROSITE; PS00065; D_2_HYDROXYACID_DH_1; 1.				
DR	PROSITE; PS00670; D_2_HYDROXYACID_DH_2; 1.				
KW	PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.				
FT	Oxidoreductase; NAD.				
NON_TER	NON_TER 1				
NN	NON_TER 1				
FE	248 248				
SEQ	SEQUENCE 248 AA; 27370 MW; 45E6ICFAFD55A0C8 CRC64;				

Query Match	24.4%	Score 520	DB 10	Length 248
Best Local Similarity	50.8%	Pred. No. 1.1e34		
Matches 101	Conservative 29	Mismatches 69	Indels 0	Gaps 0

[illegible]

Search completed: July 24, 2003, 20:48:32
Job time : 115 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:40:32 ; Search time 24 Seconds

(without alignments)
785.738 Million cell updates/sec

Title: US-09-996-008b-2

Perfect score: 2135

Sequence: 1 MAKVLCVLYDDPVDGPKTY.....YKGNATGSEBAKFKKAV 401

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwIsProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2117	99.2	400	1	FDH_PSESR
2	910.5	42.6	375	1	FDH_NEUCR
3	883.5	41.4	377	1	FDH_EMENI
4	872	40.8	384	1	FDH_ARATH
5	868	40.7	381	1	FDH_SOLTU
6	845.5	39.6	376	1	FDH_ORISA
7	843	39.5	377	1	FDH_HORVU
8	826	38.7	361	1	FDH_PICAN
9	348	16.3	525	1	SERA_METH
10	338	15.8	525	1	SERA_BACSU
11	323.5	15.2	524	1	SERA_METJA
12	314	14.7	321	1	DHG_HYPM
13	309.5	14.5	440	1	CTBP_XENLA
14	307.5	14.4	440	1	CTPI_HUMAN
15	306.5	14.4	440	1	CTPI_MOUSE
16	302.5	14.2	430	1	CTPI_RAT
17	302	14.1	437	1	CTBO_XENLA
18	287	13.9	410	1	SERA_HAEIN
19	280	13.6	527	1	SERA_ARCFU
20	287.5	13.5	533	1	SERA_RAT
21	285.5	13.4	533	1	SERA_HUMAN
22	284	13.3	315	1	YF56_HAEIN
23	281.5	13.2	528	1	SERA_MYLE
24	280	13.1	386	1	CTBP_DROME
25	277	13.0	313	1	DHG_METEX
26	276.5	13.0	445	1	CTP2_MOUSE
27	276.5	13.0	624	1	SERA_ARATH
28	275.5	12.9	445	1	CTP2_HUMAN
29	274	12.8	325	1	TKRA_BACSU
30	273.5	12.8	528	1	SERA_MYCTU
31	272	12.7	350	1	YN14_YEAST
32	269.5	12.6	485	1	SERA_MOUSE
33	264	12.4	469	1	SERA_YEAST

34	261	12.2	331	1	DDH_ZYMMO	P30799 zymomonas m
35	258	12.1	409	1	SERA_ECOLI	P08328 escherichia
36	255	11.9	336	1	PTXD_PSEST	O69054 pseudomonas
37	254.5	11.9	466	1	SERA_SCHPO	P87228 schizosacch
38	252	11.8	322	1	VANH_ENTFC	O05709 enterococci
39	252	11.8	554	1	SERA_SYNYC	P73821 synecocyst
40	248	11.6	323	1	VANH_ENTFA	O47748 enterococci
41	244	11.4	332	1	LDHD_LACDE	P26297 lactobacill
42	244	11.4	469	1	SEB3_YEAST	P40510 saccharomyc
43	241.5	11.3	334	1	YEAP_SCHPO	O14075 schizosacch
44	235	11.0	331	1	DDH_HAEIN	P44501 haemophilus
45	232	10.9	336	1	LDHD_LACHE	P30901 lactobacill

ALIGNMENTS

RESULT 1	FDH_PSESR	STANDARD	PRT	400 AA.
AC	P33160;			
DT	01-OCT-1993 (Rel. 27, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Formate dehydrogenase (EC 1.2.1.2) (NAD-dependent formate dehydrogenase) (FDH).			
DE	Pseudomonas sp. (Strain 101) (Achromobacter parvulus T1).			
OC	Bacteria; Proteobacteria.			
OX	NCBI_TaxID=33067;			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=92063805; PubMed=1954846;			
RA	Tishkov V.I., Galikin A.G., Egorov A.M.;			
RT	"NAD-dependent formate dehydrogenase of methylotrophic bacteria structure.";			
RT	dokl. Akad. Nauk SSSR 317:745-748(1991).			
RN	[2]			
RP	SEQUENCE OF 1-393.			
RX	MEDLINE=90290536; PubMed=2357236;			
RA	Popov V.O., Shumilin I.A., Ustinikova T.B., Lamzin V.S., Egorov T.A.;			
RT	"NAD-dependent formate dehydrogenase from methylotrophic bacteria Pseudomonas sp. 101. I. Amino acid sequence.";			
RL	Biorg. Khim. 16:324-335(1990).			
RN	[3]			
RP	X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).			
RX	MEDLINE=92283270; PubMed=1597184;			
RA	Lamzin V.S., Aleshin A.E., Stokopylov B.V., Yukhnevich M.G.,			
RT	Popov V.O., Harutyunyan E.H., Wilson K.S.;			
RL	"Crystal structure of NAD-dependent formate dehydrogenase.";			
RN	Eur. J. Biochem. 206:441-452(1992).			
RP	[4]			
RX	MUTAGENESIS OF CYS-255.			
RA	MEDLINE=93249485; PubMed=8484798;			
RT	Tishkov V.I., Galikin A.G., Marchenko G.N., Egorova O.A., Sheluho D.V.,			
RT	Kulakova L.B., Dementeva L.A., Egorov A.M.;			
RT	"Catalytic properties and stability of a Pseudomonas sp.101 formate dehydrogenase mutants containing Cys-255-Ser and Cys-255-Met replacements.";			
RL	Biochem. Biophys. Res. Commun. 192:976-981(1993).			
CC	-1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.			
CC	-1- SUBUNIT: Homodimer.			
CC	-1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.			
CC	PIR: J00334; J00334.			
DR	PDB: 2NAC; 26-JAN-95.			
DR	PDB: 2NAD; 26-JAN-95.			
DR	InterPro: IPR006139; 2-Hacid_DH.			
DR	InterPro: IPR006140; 2-Hacid_DH_C.			
DR	InterPro: IPR00389; 2-Hacid_DH.			
DR	Pfam: PF00286; 2-Hacid_DH_C; 1.			
DR	Pfam: PF00286; 2-Hacid_DH_C; 1.			
DR	PROSITE: PS00065; D-2-HYDROXYACID_DH_1; 1.			
DR	PROSITE: PS00670; D-2-HYDROXYACID_DH_2; 1.			

DR PROSITE; PS00671; D_2-HYDROXYACID_DH_3; 1.
 KW Oxidoreductase; NAD; 3D-structure.
 FT INIT_MET 0 146
 FT DOMAIN 1 146
 FT DOMAIN 147 333
 FT DOMAIN 334 400
 FT NP_BIND 192 226
 FT ACT_SITE 5 5
 FT ACT_SITE 284 284
 FT MUTAGEN 255 255
 FT
 FT CONFLICT 77 77
 FT CONFLICT 138 139
 FT CONFLICT 145 145
 FT CONFLICT 215 216
 FT CONFLICT 327 327
 FT STRAND 2 6
 FT STRAND 8 8
 FT TURN 12 13
 FT TURN 31 32
 FT TURN 46 47
 FT STRAND 52 52
 FT TURN 53 55
 FT HELIX 56 58
 FT HELIX 60 65
 FT TURN 66 67
 FT STRAND 69 73
 FT TURN 79 80
 FT HELIX 82 87
 FT TURN 88 89
 FT STRAND 92 96
 FT TURN 97 98
 FT STRAND 99 99
 FT STRAND 103 103
 FT HELIX 105 110
 FT TURN 112 113
 FT TURN 116 119
 FT TURN 125 126
 FT STRAND 127 127
 FT HELIX 129 134
 FT TURN 135 136
 FT STRAND 138 141
 FT TURN 143 144
 FT HELIX 147 162
 FT TURN 163 164
 FT HELIX 165 173
 FT TURN 174 175
 FT HELIX 179 183
 FT TURN 184 185
 FT TURN 190 191
 FT STRAND 193 197
 FT HELIX 201 210
 FT HELIX 211 213
 FT TURN 214 214
 FT STRAND 216 220
 FT TURN 227 233
 FT HELIX 234 234
 FT STRAND 235 237
 FT HELIX 241 244
 FT HELIX 245 247
 FT STRAND 250 253
 FT TURN 259 263
 FT STRAND 265 265
 FT HELIX 267 270
 FT TURN 271 272
 FT TURN 275 276
 FT STRAND 277 281
 FT HELIX 285 287
 FT STRAND 288 288
 FT HELIX 290 298
 FT TURN 299 300
 FT STRAND 302 307

CATALYTIC.
 COENZYME-BINDING.
 CATALYTIC.
 NAD (ADP PART).
 POTENTIAL.
 SUBSTRATE BINDING.
 C->S,M: HIGH RESISTANCE TO INACTIVATION
 BY HG(2+), INCREASED STABILITY AT 25 C.
 AND DECREASED THERMOSTABILITY AT 45 C.
 D->S (IN REF. 1).
 TV->VT (IN REF. 1).
 C->V (IN REF. 1).
 VH->HV (IN REF. 1).
 N->D (IN REF. 1).

FT TURN 317 318
 FT HELIX 320 322
 FT TURN 323 323
 FT TURN 325 326
 FT STRAND 327 327
 FT STRAND 334 336
 FT HELIX 338 357
 FT TURN 358 358
 FT HELIX 363 365
 FT STRAND 366 368
 FT STRAND 373 373
 FT TURN 375 375
 FT HELIX 376 380
 FT SO SEQUENCE 400 AA; 44005 MW; A354151D4312991F CRC64;
 Query Match 99.28; Score 2117; DB 1; Length 400;
 Best Local Similarity 99.58; Pred. No. 9.1e-172;
 Matches 398; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 AKVLCVLYDDPVDGYPKTYARDLPRKIDHYPGGQILPTPKAIDFTPGQLGVSYGELGLR 61
 DB 1 AKVLCVLYDDPVDGYPKTYARDLPRKIDHYPGGQILPTPKAIDFTPGQLGVSYGELGLR 60
 QY 62 PYLESNHGLTVYTSKDGPPDSVFRELYVDADYVISQFPWPAYLTPERIARAKNKLALTA 121
 DB 61 KYLESNGHTLVYTSKDGPPDSVFRELYVDADYVISQFPWPAYLTPERIARAKNKLALTA 120
 QY 122 GIGSDHYDLOSADIRNTVAEVTYCNSTISVAEHVMMILSLVRNYLPSHMARKGNIA 181
 DB 121 GIGSDHYDLOSADIRNTVAEVTYCNSTISVAEHVMMILSLVRNYLPSHMARKGNIA 180
 QY 182 DCVSHAYDLEAMHVTVAAGRIGLAVLRRLAPFDVHLHYTDNRRLPESVEKELNTJWHAT 241
 DB 181 DCVSHAYDLEAMHVTVAAGRIGLAVLRRLAPFDVHLHYTDNRRLPESVEKELNTJWHAT 240
 QY 242 REDMYPCDDVYTLNCPHPETEEMINDETLLKLFKRGAYIYNTARGKLCDDAVARALESG 301
 DB 241 REDMYPCDDVYTLNCPHPETEEMINDETLLKLFKRGAYIYNTARGKLCDDAVARALESG 300
 QY 302 RLAGYAGDVMPPOPAPDHPMPRTMPYNGMTPHISGTLTFOARAACTREILECFEGGRP 361
 DB 301 RLAGYAGDVMPPOPAPDHPMPRTMPYNGMTPHISGTLTFOARAACTREILECFEGGRP 360
 QY 362 IRDEYLLIVGGALAGTAGHSHSKGNATGSGEAAKFKKAV 401
 DB 361 IRDEYLLIVGGALAGTAGHSHSKGNATGSGEAAKFKKAV 400
 RESULT 2
 FDH_NEUCR STANDARD: PRT; 375 AA.
 ID 007103;
 AC 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Formate dehydrogenase (EC 1.2.1.2) (NAD-dependent formate
 dehydrogenase) (FDH).
 DE FDH OR 99H12.160.
 GN Neurospora crassa.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 OX 11
 RP SEQUENCE FROM N.A.
 RP STRAIN-74-OR23-1A / FGSC 987;
 RX MEDLINE-93285982; PubMed-8509325;
 RA Chow C.M., Rajbhandary U.U.,
 RT "Developmental regulation of the gene for formate dehydrogenase in
 J. Bacteriol. 175:3703-3709(1993).
 RL 121
 RN SEQUENCE FROM N.A.
 RC STRAIN-74-OR23-1A / FGSC 987;

RA PubMed-12655011;
 RA Manhaupt G., Montione C., Haase D., Mewes H.-W., Altmann V.,
 RA Hohelsel J.D., Fartmann B., Nakamura G., Kempken F., Mäler J.,
 RA Schulte U.;
 RT "What's in the genome of a filamentous fungus? Analysis of the
 RT Neurospora genome sequence."
 RL Nucleic Acids Res. 31:1944-1954(2003).
 CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
 CC -1- SUBUNIT: Homodimer (by similarity).
 CC -1- DEVELOPMENTAL STAGE: DEVELOPMENTALLY REGULATED. EXPRESSED ONLY
 CC DURING CONIDATION AND EARLY GERMINATION.
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 CC DEHYDROGENASES FAMILY. STRONG. TO OTHER FDH.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC -----

DR EMBL: L13964; AAA99900.1; -
 DR EMBL: AL451018; CAC18252.1; -
 DR PIR: A47117; A47117.
 DR HSSP: P33160; 2NAC.
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam: PF00389; 2-Hacid_DH; 1.
 DR Pfam: PF02826; 2-Hacid_DH.C; 1.
 DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
 KM Oxidoreductase: NAD.
 FT NP_BIND 166 201 NAD (ADP PART) (BY SIMILARITY).
 FT ACT_SITE 259 259 SUBSTRATE BINDING (BY SIMILARITY).
 FT DOMAIN 364 374 ALA-RICH.
 SQ SEQUENCE 375 AA; 40956 MW; 307395FB204968 CRC64;

Query Match 42.6%; Score 910.5; DB 1; Length 375;
 Best Local Similarity 48.4%; Pred. No. 1,4e-69;

Matches 196; Conservative 56; Mismatches 112; Indels 41; Gaps 8;

QY 1 MAKVLCVLPDVPDPRKTVARDDPKIDHYRGQILPTPKAIDFTPGQLGSGVSGEL 60
 DB 1 MKVLAIVYDGGKHS-----EVP-----ELGTFONEJGL 31
 QY 61 RPYLESNGHTLVTSDDKDPSPFERELVDADVISOFPWPAVLTPERIAKAKNLALT 120
 DB 32 RKMLEDQGHVLTTCDDKNGENSTFDEKLEDAEIIITTFPHGYLAERLAKKTKLAVT 91
 QY 121 AGISGDHVDLOSADRN--VVAEYTYGNSISVAEHVYMMILSVRNLPSHEMARKGW 178
 DB 92 AGISGDHVDLNAKNTNGITVAEYTGNSVSAHVMTLTVLRNVPAPHEQIQEGSW 151
 QY 179 NIADCVSHAYDLEAMHGTVAAGRIGLAVLRRLAPFDV-HLYTDRHRLPESVEKELNT 237
 DB 152 DVAEAKNEPDEGVVGTGVRIGERYLRKLPFDDKELLYIYQPLSAKEKEIGCR 211
 QY 238 WHATREDMYPVCDVVTINCPHPETEMINDETLKFRGAVIVNTARGKLCDDRAVARA 297
 DB 212 RVADLEEMLAQCDVVTINCPHPETEMINDETLKFRGAVIVNTARGKLCDDRAVARA 271
 QY 298 LESGLAGVAGVMPROPAPKDPHPR--TMPY--NGMTPHISGTTLRQARYAAGTREI 352
 DB 272 LKSGLRGVDGVMPPAPDPHPLRYAKNPFGGNAVPHNSGSLDAOKRYAGTAKI 331
 QY 353 LECFEGR-PIRDELYIVOGGALAGTGAHSYSGKATGSGSEPAK 396
 DB 332 IESYLSGKHDRPEDLIYGGDYA---TKSIGERERAAAAAAAK 373

RESULT 3

FDH_EMENT
 ID FDH_EMENT STANDARD; PRT; 377 AA.

AC 003134;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable formate dehydrogenase (EC 1.2.1.2) (NAD-dependent formate
 DE dehydrogenase) (FDH).
 GN ACIA.
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Emericella.
 OX NCBI_TaxID=162425;

RN [1]
 RP MEDLINE-93101140; PubMed-1465107;
 RA Saleeba J.A., Cobbett C.S., Hynes M.J.;
 RT "Characterization of the anda-regulated acia gene of Aspergillus
 RT nidulans".
 RL Mol. Gen. Genet. 235:349-358(1992).
 CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
 CC -1- SUBUNIT: Homodimer (by similarity).
 CC -1- INDUCTION: ACETATE INDUCTION MEDIATED BY AMDA REGULATORY GENE.
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 CC DEHYDROGENASES FAMILY. STRONG. TO OTHER FDH.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC -----

DR EMBL: 211612; CAA7687.1; -
 DR PIR: S30088; S30088.
 DR HSSP: P33160; 2NAC.
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam: PF00389; 2-Hacid_DH; 1.
 DR Pfam: PF02826; 2-Hacid_DH.C; 1.
 DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; FALSE_NEG.
 KM Oxidoreductase: NAD.
 FT NP_BIND 161 191 NAD (ADP PART) (BY SIMILARITY).
 FT ACT_SITE 271 271 SUBSTRATE BINDING (BY SIMILARITY).
 FT DOMAIN 377 AA; 41531 MW; 52195FEBCC931CE CRC64;

Query Match 41.4%; Score 883.5; DB 1; Length 377;
 Best Local Similarity 52.1%; Pred. No. 2.7e-67;

Matches 185; Conservative 48; Mismatches 95; Indels 27; Gaps 6;

QY 47 PGQLGYSVSGELRLPYLESNGHTLVTSDDKDPSPFERELVDADVISOFPWPAVLT 106
 DB 14 PG-LIGTLENELGIRKMWIEQGHVLTTSDDKNGENSTFDEKLEDAEIIITTFPHGYLTA 72
 QY 107 ERIAKAKNLALTGIGSDHVDLOSADRN--VVAEYTYGNSISVAEHVYMMILSVR 164
 DB 73 ERIAKAKNLKLAVTGIGSDHVDLDAKNTNGITVAEYTGNSVSAHVMTLTVLR 132
 QY 165 NYLPSHEMARKGWNIADCVSHAYDLEAMHGTVAAGRIGLAVLRRLAPFDV-HLYTDR 223
 DB 133 NFVPAHDQIRNDNMVAAYAKNEFDENKVVGTGVRIGERYLRKLPFDDKELLYYD 192
 QY 224 HRLPESVEKELNLTGHAATREDMYPVCDVVTINCPHPETEMINDETLK----- 272
 DB 193 QPLREVEKEIGARVDSLEEVYQCDVVTINCPHPETEMINDETLK----- 252
 QY 273 -----LRRKAVIYNTARGKLCDDRAVARALESGRLAGVAGVMPROPAPKDPHPR 326
 DB 253 LIIPMLMYHKGSLVNTARGAIVKEDVAELKSGHLRNGYGDVMPROPAPKDPHPR 312

QY 327 Y-----NGMPHISGTTTAAQARYAGTRETILECFEGGR-PIRDEYLIVGGGALA 375
 Db 313 HPMGGGATVPMGSGTSLAQIRYANGTKALIDSYSGRFDPYQODLIVHGDDYA 367

RESULT 4
 FDH_ARATH STANDARD; PRT; 384 AA.
 AC Q957E4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Formate dehydrogenase, mitochondrial precursor (EC 1.2.1.2) (NAD-dependent formate dehydrogenase) (FDH).
 GN FDH1 OR FDH OR AT5G14780 OR T9L3_80.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eustosids; II: Brassicales; Brassicaceae; Arabidopsis.
 CX NCBI_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nishikawa T., Fukusaki E., Kobayashi A.;
 RT "Formate dehydrogenase cDNA from Arabidopsis thaliana."
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Li R., Bonham-Smith P.C., King J.;
 RT "Arabidopsis thaliana mRNA for NAD-dependent formate dehydrogenase 1."
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Markwell J., Osterman J.C., Olson B.J., Skavdahl M., Ramberg H.,
 RA Hermann M.C.;
 RT "Induction of leaf formate dehydrogenase by one-carbon metabolites."
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016721; PubMed=11130714;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
 RA Nakazaki N., Nartuo K., Okumura S., Shitipo S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
 RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
 RA Stoneking T., Pepin K., Splieth J., Sekhon M., Armstrong J., Becker M.,
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
 RA Du H., Edwards J., Elyman J., Haakensen B., Lamar E., Latreille P.,
 RA Leonard S., Meyer R., Mullaney E., Ozersky P., Riley A., Stromatt C.,
 RA Wagner-McPherson C., Wolman A., Yeakum M., Bell M., Dedina N.,
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vill D., Baker J.,
 RA Kirchhoff K., Toth K., King L., Bahret A., Miller B., Marra M.,
 RA Mattiensen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Volckert G., Wambolt R., Duesterhoft A., Stekema W., Pohl T.,
 RA Enlian K.-D., Terryn N., Hartley N., Bent E., Johnson S.,
 RA Langham S.-A., McCullagh B., Robben J., Grymoprez B., Zimmermann W.,
 RA Ransperger U., Wedler H., Balke K., Wedler E., Peters S.,
 RA van Staveren M., Dirks W., Woolman P., Klein Lankhorst R.,
 RA Weltzenegger T., Bothe G., Rose M., Haut J., Berneriser S., Hempel S.,
 RA Feldpusen M., Lambers S., Villarroel R., Giesen J., Adiles W.,
 RA Berts O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.;
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis thaliana."
 RL Nature 408:823-826(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the

RT SSP consortium (Salk/Stanford/PGEC).";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 CC DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB023897; BAA8683.1; -
 DR EMBL; AF208028; AAF19435.1; -
 DR EMBL; AF208029; AAF19436.1; -
 DR EMBL; AF217195; AAF67100.1; -
 DR EMBL; AL391149; CAC01877.1; -
 DR EMBL; AY054285; AAL06944.1; -
 DR EMBL; AY039609; AAK62664.1; -
 DR EMBL; AY081734; AAL87387.1; -
 DR PIR; T51423; T51423.
 DR HSSP; P33160; 2NAD.
 DR InterPro; IPR006139; 2-Hacid_DH.
 DR InterPro; IPR006140; 2-Hacid_DH.C.
 DR Pfam; PF00389; 2-Hacid_DH.1.
 DR Pfam; PF02825; 2-Hacid_DH.C.1.
 DR PROSITE; PS00605; D_2-HYDROXYACID_DH.1; 1.
 DR PROSITE; PS00670; D_2-HYDROXYACID_DH.2; 1.
 DR PROSITE; PS00671; D_2-HYDROXYACID_DH.3; FALSE NEG.
 KW Oxidoreductase; NAD; Mitochondrion; Transist peptidase.
 FT TRANSIT 1 27 MITOCHONDRION (BY SIMILARITY).
 FT CHAIN 28 384 MITOCHONDRION (BY SIMILARITY).
 FT NP_BIND 198 232 NAD (ADP PART) (BY SIMILARITY).
 FT ACT_SITE 290 290 SUBSTRATE BINDING (BY SIMILARITY).
 FT SEQUENCE 384 AA; 42409 MW; A12BA423019D862B CRC64;
 SO Query Match 40.8%; Score 872; DB 1; Length 384;
 Best Local Similarity 52.9%; Pred. No. 2.6e-66;
 Matches 172; Conservative 44; Mismatches 109; Indels 0; Gaps 0;

QY 51 LGSVSGELGRPYESNGHTLVTSXKDGPDSEFERELVDADVISOPEMPAYITPPIRIA 110
 Db 56 LGCVENLGIHDMVDSOGHOYIVYDKEGPCCELEKHIPDLHYLSTPFPAYTAIRIK 115

QY 111 KAKULKALTPAGTSDHVDLOSATIDRNVYVAEVTGNSISVAEHVMMITSLVNTYPSH 170
 Db 116 KAKULKALLTRIGISDHDIDQAAAGLVAEVTGNSVSAEDELKRLILIMRNFPVPGY 175

QY 171 EMARKGMNTADCVSHAYDLEAMHGVTVAGRTGLAVLRRLAPDVHLHYDRRLPEVY 230
 Db 176 NQVYKGMNVAGIAYRAYVDELGKTIYGVAGRIKLLQRIKPGCNLLHDRQMAPEL 235

QY 231 EKEILNTWATREDMYPVCVYVTLNCPJLHPETEMINDETLKFKRQAYIVNTARGKLC 290
 Db 236 EKEKGAFVEDLNMELKPCQIVYINMPLTETRCMFNKEILGKLKGCVLIVNNARGAIME 295

QY 291 RDAVARLESGRLAGYGVDFWDFOPAPKDPHWRMPNMGMPHISGTTTAAQARYAGTR 350
 Db 296 RQAVVADVESHGIGSYGDWDFOPAPKDPHWRMPNMGMPHISGTTTAAQARYAGTR 355

QY 351 EILECFEGRPIRDEYLIVGGGALA 375
 Db 356 DMLERYKGGDFPENTENTIVDGEIA 380

RESULT 5
 FDH_SOLTU STANDARD; PRT; 381 AA.
 AC 007511: 09ZR28;
 DT 01-FEB-1995 (Rel. 31, Created)

```

DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Formate dehydrogenase, mitochondrial precursor (EC 1.2.1.2) (NAD-
DE dependent formate dehydrogenase) (FDH).
GN FDH1.
OC Solanum tuberosum (Potato); Streptophyta; Embryophyta; Tracheophyta;
OC Eukaryota; Viridiplantae;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; solanales; Solanaceae; Solanum.
OX NCBI_Taxid=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. BF15;
RA Hourton-Cabassa C., Ambard-Bretteville F., Remy R.,
RA Colas des Francs-Small C.;
RT "Evidence for multiple copies of formate dehydrogenase genes in
RT plants: Isolation of three potato fdh genes fdh1, fdh2 and fdh3."
RL (in) Plant Gene Register FGR98-102.
RN [2]
RP SEQUENCE OF 3-381 FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN-CV. BF15; TISSUE=Tuber;
RX MEDLINE=94105343; PubMed=8278546;
RA Colas des Francs-Small C., Ambard-Bretteville F., Small I.D.,
RA Remy R.;
RT "Identification of a major soluble protein in mitochondria from
RT nonphotosynthetic tissues as NAD-dependent formate dehydrogenase."
RL Plant Physiol. 102:1171-1177(1993).
RN [3]
RP REVISIONS TO N-TERMINUS.
RC STRAIN-CV. BF15; TISSUE=Tuber;
RA Colas des Francs-Small C.C.;
RL submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 26-54.
RC STRAIN-CV. BF15; TISSUE=Tuber;
RA Colas des Francs-Small C., Ambard-Bretteville F., Darpas A.,
RA Sallantin M., Ruet J.-C., Pernollet J.-C., Remy R.;
RT "Variation of the polypeptide composition of mitochondria isolated
RT from different potato tissues."
RL Plant Physiol. 98:273-278(1992).
RN [5]
RP CHARACTERIZATION.
RC PubMed=9490763;
RA Hourton-Cabassa C., Ambard-Bretteville F., Moreau F.,
RA Daye de Virville J., Remy R., Colas des Francs-Small C.;
RT "Stress induction of mitochondrial formate dehydrogenase in potato
RT leaves."
RL Plant Physiol. 116:627-635(1998).
CC -1- FUNCTION: Involved in formate-dependent oxygen uptake coupled to
CC ATP synthesis.
CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- TISSUE SPECIFICITY: Found at high levels in developing tubers, at
CC intermediate level in stems, veins, stolons, and stemens,
CC and at low level in leaves and roots.
CC -1- INDUCTION: Induced very rapidly by wounding, and slower by
CC darkness, chilling, drought, hypoxia, and treatments with formate,
CC abscisic acid, serine, sarcosine, pyruvate, acetate, ethanol or
CC methanol.
CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.
CC -1- CAUTION: There are two other putative pseudogenes, fdh2 and fdh3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on ways
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcements
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z99991; CAB17080.1; -.

```

DR EMBL; Z699992..: NOT ANNOTATED_CDS.
DR EMBL; Z21493; CAA79702.2; ..
DR PIR; JQ2272; JQ2272.
DR HSSP; P33160; 2NAD.
DR InterPro; IPR006139; 2-Hacid_DH.
DR InterPro; IPR006140; 2-Hacid_DH.C.
DR Pfam; PF00389; 2-Hacid_DH.1.
DR Pfam; PF02826; 2-Hacid_DH.C.1.
DR PROSITE; PS00065; D_2-HYDROXYACID_DH.1; 1.
DR PROSITE; PS00670; D_2-HYDROXYACID_DH.2; 1.
DR PROSITE; PS00671; D_2-HYDROXYACID_DH.3; 1.
DR OXidoreductase; NAD; Mitochondrion; Transit peptide.
KW TRANSIT
FT CHAIN 1 25 MITOCHONDRION.
FT NP_BIND 26 381 FORMATE DEHYDROGENASE.
FT ACT_SITE 195 229 NAD (AP PART) (BY SIMILARITY).
FT SEQUENCE 287 287 SUBSTRATE BINDING (BT SIMILARITY).
SO QUERY_MATCH 361 AA; 42038 MW; 9200J3IAC5A3A00E CRC64;

Query Match Best Local Similarity 40.7%; Score 868; DB 1; Length 381;
Matches 175; Conservative 48; Mismatches 117; Indels 12; Gaps. 2.

QY 35 QILETPRAI-----DPTGQLGSVSGELGLRPYLENSGHTLVTSKDGPSV 83
Db 27 QASGPRIYGVFKANEYAMNP-NFGCAENALGIEMWESKHQGYIVPDKGPCPE 85
QY 84 FERLEADVADVIISOPMPAYLTPIRIAKAKMKLALTAGTISDHVDLOSAIDRNTVAEV 143
Db 86 LEKIIPDLHLVLSIFPHRAYTAERIKAKMLQLLTGTIGSDHVDLKAAAAGTLVAEV 145
QY 144 TYCNSISVAEHVMMLSLVRNYLPSEHMARKGMNIADCVSNAVDEAMHVGAAGRI 203
Db 146 TGSNTVAEDELKRILLVNFPLPGHQVINGEWNVAAIAHRPADEGKTGVGAGRI 205
QY 204 GLAVLRRLARPDAVLHTDRRLDESVEKEINTFMHARRDMRYPCDVVTINCPHPETE 263
Db 206 GRLLQRKPKNCLMHLDRCLKMSSELNQIGAKFEEDLMKLSKCDIVVINTPLETKK 265
QY 264 HMINDETLFLEKRGAVLYNTARGICDRDAYARALSRLTAGVADVPFPAPKDPWR 323
Db 266 GMFKERIATKLKKGVLLVNNARGAIMDTQAIVDACNSGHIGSDDVYPPAPKDPWR 325
QY 324 TMPLNGMTPHISGTTLNQAARYAAGTRILECFEPGRIRIDREYLIVOGALA 375
Db 326 YMPQAMTPHISGTTIDAQLRYAAGTYMDLDRYKGEDEPFAENVYKGDGLA 377

RESULT 6
FDH_ORYA
ID FDH_ORYA STANDARD: PRI: 376 AA.
AC Q95XP2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Formate dehydrogenase, mitochondrial precursor (EC 1.2.1.2) (NAD-dependent formate dehydrogenase) (FDH).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzeae; Oryza.
OX NCBI_TaxId=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Shiraishi T., Fukusaki E., Kobayashi A.;
RT "NAD-dependent formate dehydrogenase.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
CC -1- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

OX NCBI_TaxID=4905;
 RP SEQUENCE FROM N.A.
 RA Hollenberg C.P., Janowicz Z.;
 RT "DNA molecules coding for FMDH control regions and structured gene for
 RT a protein having FMDH-activity and their uses.";
 RL Patent number EP0299108, 18-JAN-1989.
 CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) -> CO(2) + NADH.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 CC DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: A06214; CAA00531.1; -
 DR HSSP: P33160; 2NAD
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam: PF00389; 2-Hacid_DH; 1.
 DR Pfam: PF02826; 2-Hacid_DH.C; 1.
 DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
 DR Oxidoreductase; NAD.
 FT INT_MET 0 PROBABLE.
 FT NP_BIND 164 199 NAD (ADP PART) (BY SIMILARITY).
 FT ACT_SITE 257 257 SUBSTRATE BINDING (BY SIMILARITY).
 SQ SEQUENCE 361 AA; 39779 MW; 0FC001366F9E47B9 CMC64;

Query Match 38.7%; Score 826; DB 1; Length 361;
 Best Local Similarity 45.1%; Pred. No. 1.9e-62;
 Matches 171; Conservative 63; Mismatches 107; Indels 38; Gaps 6;
 OY 3 KVLGVLDPPDYKPKYARDLPKIDHPGGQLPPRKALDTPFGQLGSGSELGRP 62
 DB 1 KVLGVLDY-----AGKHADE-----ERLYGTENALGIRD 31
 OY 63 YLESNGHTLVTSKDPDSYFERELVDADAVISQFPWPAYLTPERTAKAKNLKIALTAG 122
 DB 32 WLEKQGHVYVTSKDEQNSVLEKINISDADVIITSPHPATITKERIDKAKLLVYAG 91
 OY 123 IGSDDVLDQ--SAIDRNVTVAEVTYCNISIVAEHVVMKILSVRNLYPSHEMARKGNNI 180
 DB 92 VGSDDHLDYINQSGRDISVLETVGSNVVSAEHHVMTMLVLRNFPVPAHQIISGGNNV 151
 OY 181 ADCSHAYVDEAMNVTVAAGRIGLAVLRRLAPD-VHLATDTHRRLPESVEKELNLTWH 239
 DB 152 AEIAKSDPDLGKVIATIGARIGYRVLERLVAENPREKLLYYDQSLKEEKEVGARRV 211
 OY 240 ATREDMYPVCDVYVTLNCPLEHPTHEMINDETLLKLFKGAIVYNTARGKLCORDVAARLE 299
 DB 212 HDIKELVAQADIVYINCPLEHAGSGLVNAELIKFKKGAIVYNTARGAICVAEDVAAAVK 271
 OY 300 SGRLAGYAGDVWEPQAPKDPKHPMTPEY-----NGMPTHSIGTLTAQARAAGTRETLE 354
 DB 272 SGLRGYGGDVWEPQAPKDPKHPMSMANKYAGANAMPYHSGSVIADQVRYAAGTKNITLE 331
 OY 355 CFEGR-PIRDEYIIVGG 372
 DB 332 SFTQKEDYRPDIITLLNG 350

RESULT 9
 ID SERA_METH STANDARD; PRT; 525 AA.
 AC 027051;
 DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) (PGDH).
 GN SERA OR MTH970.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RP SEQUENCE FROM N.A.
 RC STRAIN=Delta H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert R.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
 RA Spadafora R., Vitcare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jivani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA Mchougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT deltaH: functional analysis and comparative genomics.";
 RL J. Bacteriol. 179:7135-7155(1997).
 CC -1- CATALYTIC ACTIVITY: 3-phosphoglycerate + NAD(+) -> 3-
 CC phosphohydroxypyruvate + NADH.
 CC -1- PATHWAY: Serine biosynthesis; first step.
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 CC DEHYDROGENASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE000870; AAB85466.1; -
 DR PIR: H69229; H69229.
 DR HSSP: P08328; 1PSD.
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR InterPro: IPR002912; ACT.
 DR InterPro: IPR006236; PGDH.
 DR Pfam: PF00389; 2-Hacid_DH; 1.
 DR Pfam: PF02826; 2-Hacid_DH.C; 1.
 DR Pfam: PF01842; ACT; 1.
 DR TIGRfams: TIGR01327; PGDH; 1.
 DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
 KW Serine biosynthesis; Oxidoreductase; NAD: Complete proteome.
 FT ACT_SITE 232 232 SUBSTRATE BINDING (BY SIMILARITY).
 FT ACT_SITE 261 261 BY SIMILARITY.
 FT ACT_SITE 279 279 BY SIMILARITY.
 SQ SEQUENCE 525 AA; 56990 MW; DDD63CF8F812B1C5 CMC64;

Query Match 16.3%; Score 348; DB 1; Length 525;
 Best Local Similarity 31.7%; Pred. No. 9.9e-22;
 Matches 97; Conservative 54; Mismatches 131; Indels 24; Gaps 6;
 OY 72 VVTSDDKQDPSYFERELVDADAVISQFPWPAYL-----TPERIAKAK 113
 DB 6 VLINDSINEKISISELEV-AEYVYVNTTTPPELLDAIKDPALVYRSKTKYTRVEIAP 64
 OY 114 NKLALTAGIGSDHVDLSAIDRNVTVAEVTYCNISIVAEHVVMKILSVRNLYPSHEMA 173
 DB 65 RLKTIIRAGVGVNDVYAAATDRGIMVINAPESTISITVAEHSIGLMALARKIAIDRSV 124
 OY 174 RKGGMNADCVSHAYVDEAMNVTVAAGRIGLAVLRRLAPDVLHATDTHRRLPESVEKE 233
 DB 125 KEGWENKRFM--GIEINGKTLGIIGMRIGSQVYVTKRKGMDIMYDPIKSEAE-E 181
 OY 234 LNTWHAATREDMYPVCDVYVTLNCPLEHPTHEMINDETLLKLFKGAIVYNTARGKLCORDA 293

```

Db      182 MCVT-VTDLFTLLRESIDIVTHVLPETRHLLISEDFKLMKDAFIYVNCARGIIDEBA 240
Oy      294 VARAESGRAGYAGDVWFPQAPKDHPRMTNPGTTPHISGTLTLAOCARYAGTBEIL 353
Db      241 LYRAKGEETAGALDVFEEBP-PESGPLLELENVLTPLHIGASTSEORDAITIVANEI 299
Oy      354 ECFEEG 359
Db      300 KTFEFG 305

RESULT 10
SERB_BACSU STANDARD; PRT; 525 AA.
ID SERB_BACSU 032011;
AC P3136; 032011;
DT 01-FEB-1994 (Rel. 28, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) (PGDH).
GN SERA.
OS Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OC NCBI_TaxID=1423;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / Marburg;
RX MEDLINE=96349105; PubMed=8760912;
RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
RA Serror P.;
RT "sequence analysis of the Bacillus subtilis chromosome region between
RL the sera and kds loci cloned in a yeast artificial chromosome."
RL Microbiology 142:2005-2016(1996).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertolo M.G., Bessieres P., Bolotin A., Borcher S.,
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton J.F., Cummings N.J., Daniel R.A.,
RA Denliot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fedre C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Gallizi A., Galleron N.,
RA Gilm S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klier-Bleichard M., Klein C.,
RA Kobayashi Y., Koelter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S., Nodack M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Paro V., Pohl T.M., Portetelle D., Portwollik S., Prescott A.M.,
RA Prescan E., Pujic P., Punelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadie Y.,
RA Sato T., Scanlon E., Schleich S., Schroeter R., Scofield F.,
RA Sekiuchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Solido B.,
RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Takashi A., Tanaka T., Terstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassartoli A.,
RA Viart A., Wambut R., Wedler E., Wedler H., Weitzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
[3]
RP SEQUENCE OF 107-525 FROM N.A.
RC STRAIN=168 / Marburg;
RX MEDLINE=95020538; PubMed=7934829;
RA Sorokin A.V., Zumstein E., Azevedo V., Ehrlich S.D., Serror P.;
```

```

RT      "The organization of the Bacillus subtilis 168 chromosome region
RT      between the spoVA and sera genetic loci, based on sequence data.";
RL      Mol. Microbiol. 10:385-395(1993).
CC      -1- CATALYTIC ACTIVITY: 3-phosphoglycerate + NAD(+) -> 3-
CC      phosphohydroxypyruvate + NADH.
CC      -1- ENZYME REGULATION: IN BACTERIA DISPLAYS FEEDBACK INHIBITION BY
CC      L-SERINE.
CC      -1- PATHWAY: Serine biosynthesis; first step.
CC      -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC      DEHYDROGENASES FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; L47648; AAC83943.1; -
DR      EMBL; 299116; CAB14239.1; -
DR      EMBL; L09228; AAA67502.1; -
DR      PIR; C69705; C69705.
DR      HSSP; P33160; ZNAD.
DR      Subtilist; BG10509; sera.
DR      Interpro; IPR006139; 2-Hacid_DH.
DR      Interpro; IPR006140; 2-Hacid_DH_C.
DR      Interpro; IPR002912; ACT.
DR      Interpro; IPR006236; PGDH.
DR      Pfam; PF00388; 2-Hacid_DH_1.
DR      Pfam; PF02826; 2-Hacid_DH_C; 1.
DR      Pfam; PF01842; ACT; 1.
DR      TIGRFAMS; TIGR01327; PGDH; 1.
DR      PROSITE; PS00065; D_2-HYDROXYACID_DH_1; 1.
DR      PROSITE; PS00670; D_2-HYDROXYACID_DH_2; 1.
DR      PROSITE; PS00671; D_2-HYDROXYACID_DH_3; 1.
DR      Serine biosynthesis; Oxidoreductase; NAD; Complete proteome..
KW      NP_BIND 238 267 NAD (BY SIMILARITY).
FT      ACT_SITE 229 229 SUBSTRATE BINDING (BY SIMILARITY).
FT      ACT_SITE 258 258 BY SIMILARITY.
FT      ACT_SITE 276 276 BY SIMILARITY.
SQ      SEQUENCE 525 AA; 57115 MW; B3D054487DD44FF CRC64;

Query Match 15.88; Score 338; DB 1; Length 525;
Best Local Similarity 31.28; Pred. No. 6.9e-21;
Matches 102; Conservative 54; Mismatches 131; Indels 40; Gaps 9;

Oy      59 GLRPYLENSGHTLVTSKDKGPSVFERELVDADVISOFPWRYLTPERIAAKNIKLA 118
Db      14 GLQPLESDPFIETVQKNVADAED--ELATFDALVRS---ATKVTEDLFNMTSLKIV 66
Oy      119 LTFAGIGSDHVDLSAIDRNTVAEVTCNSISVAEHVMMILSVRY-----LPSHEW 172
Db      67 GRAGVGNDIDIDEATHGIVYINAPNGNTISRAEHFAMISLMRIIPANTSVSKREW 126
Oy      173 ARKGWNIADCVSHAY--DLAMHGTVAAGRIGLAVALRRLAPFDVHLHYTRHRLPES 229
Db      127 NRT-----AYGSELVYGTGLTIGVIGRIGSEIAORRGAFCWTVAHFDPF-LTEE 174
Oy      230 VEKEINTWHTATREDMPVCDVYTLNCPLEPHEHMTNDETLKFRGATVYVTAHGKLC 289
Db      175 RAKKIGN-SRTEEVLESADITVTPRLKKEKGLINKETKAGVRLINCARGIT 233
Oy      290 DRDAVARALSGLAGYAGDVWFPQAPKDHPRMTNPGTTPHISGTLTLAOCARYAGT 349
Db      234 DEAALEALENGHVAAGALDVFEEBP-PVDNKLVDHPLVATPHLGASTSEADLNVAAGV 292
Oy      350 RETLECFEGRP-----RDEX 366
Db      293 SEEVLFQAKGLPYMSAINLPAMTKDER 319

RESULT 11
```

```

SER# MET#  ID SER# MET#  STANDARD:  PRI:  524 AA.
AC Q58424;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) (PGDH).
CN SER# OR MJ1018.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID:2190;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RA But C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Kurland G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glöck A.,
RA Scott J.L., Geoghegan N.S.M., Wadman J.F., Fulhman J.L., Nguyen D.,
RA Ueberlack T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klank H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -1- CATALYTIC ACTIVITY: 3-phosphoglycerate + NAD(+) = 3-
CC phosphohydroxypyruvate + NADH.
CC -1- PATHWAY: Serine biosynthesis; first step.
CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC DEHYDROGENASES FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67544; AAB9920.1; -.
DR PIR: A64427; A64427.
DR HSSP: P08328; 1P8D.
DR TIGR: MJ1018; -.
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR InterPro: IPR002912; ACT.
DR InterPro: IPR006236; PGDH.
DR Pfam: PF00389; 2-Hacid_DH; 1.
DR Pfam: PF02826; 2-Hacid_DH.C; 1.
DR Pfam: PF01842; ACT; 1.
DR TIGRFAMs: TIGR01327; PGDH; 1.
DR PROSITE: PS0065; D_2-HYDROXYACID_DH.1; 1.
DR PROSITE: PS00670; D_2-HYDROXYACID_DH.2; 1.
DR PROSITE: PS00671; D_2-HYDROXYACID_DH.3; 1.
FT FT Serine biosynthesis; Oxidoreductase; NAD; Complete proteome.
FT NP_BIND 240 269 NAD (BY SIMILARITY).
FT ACT_SITE 231 231 SUBSTRATE BINDING (BY SIMILARITY).
FT ACT_SITE 260 260 BY SIMILARITY.
FT ACT_SITE 278 278 BY SIMILARITY.
SO SEQUENCE 524 AA; 56923 MW; 5C5561BB6C10793 CRC64;

Query Match 15.2%; Score 323.5; DB 1; Length 524;
Best Local Similarity 32.1%; Pred. No. 1.2e-19;
Matches 96; Conservative 58; Mismatches 108; Indels 37; Gaps 10;

64 LENSNGTHLYVTSKRDGPDSPVERELY---DADVYISQEPMPAYLTPEIRAKAKNLKAL 119
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 18 LEEVGEVEVAT-----GLTKRELEIKIKADADLVNRS--GTRVTRDVIKAEKLVIG 68
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
120 TAGGSGHVDLQSAIDRNVYVAEVTYGCNSISVAEHVYMMILSYRNY-----LPSHMA 173
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11

```

ID	NAME	DESCRIPTION	SEQUENCE
D6	69	RACVGDNDIDVEATKTKGIIIVNADPDAISSIVAEITLNMGLMLAARNIPQANSLKREMD	128
QY	174	RKCGMNIADCVSHAYDLLEAMHVGVAAGRIGLAVLRRLAPFDVHLHTDRHRLPESVEKE	233
D6	129	RRFRKKI-----ELYKTLGTIGVLIGGOOVVRAKAKAGMNIIGDYP-IPKEVAES	179
QY	234	LNLTWATREDMYRVC---DVTYTLNCPRLPEREHHINDETLLKLRKRAYIYNTPARGLCD	250
D6	180	MGV---ELVDDINELCKRADFTTLVHPRLPKTRHILGREGQIALKNKAIIVNCARGGLID	236
QY	291	RDVAVARALESGRIAGVDWFPQAPKDPHPRTPMYGMPHISGTLTFAQARYAGT	349
D6	237	EKLRYEALKEGKIRAAALDVEEER-PKDNPLTLTIDNYIGTFPHOGAS--TEEAQKAAGT	292
RESULT 12			
ID	DHG_YHPME	STANDARD:	PRT: 321 AA.
AC	P36234:	01-JUN-1994 (Rel. 29, Created)	
DT	01-JUN-1994 (Rel. 29, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Glucuronate dehydrogenase (EC 1.1.1.29) (NADH-dependent hydroxypyruvate reductase) (HPR) (GDH) (Hydroxypyruvate dehydrogenase) (Glyoxylate reductase).		
OS	Hyphomicrobium methylotrophicum.		
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;		
OC	Hyphomicrobiaceae; Hyphomicrobium.		
OX	NCBI_Taxid=84;		
RN	[1]		
RP	X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).		
RX	MEDLINE=94166078; Pubmed=8120891;		
RA	Goldberg J.D., Yoshida T., Brick P.;		
RT	"Crystal structure of a NAD-dependent D-glycerate dehydrogenase at		
RT	2.4-A resolution.";		
RL	J. Mol. Biol. 236:1123-1140(1994).		
RN	[2]		
RP	CHARACTERIZATION.		
RP	STRAIN=KM146 / GM2;		
RC	MEDLINE=90306024; Pubmed=2114287;		
RX	Izumai Y., Yoshida T., Kanzaki H., Toki S.-I., Miyazaki S.S.,		
RA	Yanada H.;		
RT	"Purification and characterization of hydroxypyruvate reductase from		
RT	a serine-producing methylotroph, Hyphomicrobium methylotrophicum GM2.";		
RL	Eur. J. Biochem. 190:279-284(1990).		
CC	-1- FUNCTION: ACTIVE ON HYDROXYPYRUVATE AND GLYOXYLATE. OPTIMUM PH IS		
CC	6.8 AND TEMPERATURE IS 45 DEGREES CELSIUS.		
CC	-1- CATALYTIC ACTIVITY: (R)-glycerate + NAD(+) = hydroxypyruvate +		
CC	NADH.		
CC	-1- PATHWAY: Serine pathway.		
CC	-1- SUBUNIT: Homodimer.		
CC	-1- INDUCTION: By methanol.		
CC	-1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID		
CC	DEHYDROGENASES FAMILY.		
CC	PDB: 1GDH; 3I-JAN-94.		
DR	InterPro: IPR006139; 2-Hacid_DH.		
DR	InterPro: IPR006140; 2-Hacid_DH_C.		
DR	Pfam: PF00389; 2-Hacid_DH; 1.		
DR	Pfam: PF02826; 2-Hacid_DH_C; 1.		
DR	PROSITE: PS00065; D_2-HYDROXYACID_DH_1; 1.		
DR	PROSITE: PS00670; D_2-HYDROXYACID_DH_2; 1.		
DR	PROSITE: PS00671; D_2-HYDROXYACID_DH_3; 1.		
KW	Oxidoreductase; NAD; 3D-structure.		
FT	INIT MET	0	
FT	DOMAIN	1	99
FT	DOMAIN	100	290
FT	DOMAIN	291	321
FT	NP_BIND	149	177
FT	ACT_SITE	240	240
FT	ACT_SITE	269	269
FT	ACT_SITE	287	287
FT	STRAND	4	7
FT	HELEX	13	20


```

Db      192 RAKTFGFEVNFYEDPY-LSDGIERALGLQRIYSTLDLFLHSCVTLHGLGNHNHNLNDE 250
Qy      270 TLKLFKRGATVNTARCKLDRDVAVALRESGLAGVDFPOP-APKDHPRTPMPYN 328
Db      251 TIKOMRGAGFLVNTARGSLVDKALQALKEGRIGRGAALDVHESEPSFGQPLKDAVNL 310
Qy      329 GMTPH 333
Db      311 ICTPH 315

RESULT 14
CTPL_HUMAN
ID      CTPL_HUMAN          STANDARD:      PRT:      440 AA.
AC      Q13363;
DT      15-JUL-1998 (Rel. 36, Created)
DT      15-JUL-1999 (Rel. 38, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      C-terminal binding protein 1 (CTBP1).
GN      CTBP1 OR CTBP.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A., AND SEQUENCE OF 98-108.
RC      TISSUE-B-cell, and Cervical carcinoma;
RX      MEDLINE=96068636; PubMed=7479821;
RA      Schaeper U., Boyd J.M., Verma S., Uhlmann E., Subramanian T.,
RA      Chinnadural G.;
RT      "Molecular cloning and characterization of a cellular phosphoprotein
RT      that interacts with a conserved C-terminal domain of adenovirus E1A
RT      involved in negative modulation of oncogenic transformation.";
RT      Proc. Natl. Acad. Sci. U.S.A. 92:10467-10471(1995).
RN      [2]
RP      SEQUENCE FROM N.A., AND REVISIONS.
RX      MEDLINE=99078014; PubMed=9858600;
RA      Sewalt R.G.A.B., Gunster M.J., van der Vliet J., Sattijn D.P.E.,
RA      Otte A.P.;
RT      "C-terminal binding protein is a transcriptional repressor that
RT      interacts with a specific class of vertebrate polycomb proteins.";
RT      Mol. Cell. Biol. 19:777-787(1999).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      TISSUE-Brain;
RX      MEDLINE=22386257; PubMed=12477932;
RA      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Rata S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Bulik S.W.,
RA      Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA      Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length
RT      human and mouse cDNA sequences.";
RT      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN      [4]
RP      CHARACTERIZATION.
RX      MEDLINE=93178421; PubMed=8440238;
RA      Boyd J.M., Subramanian T., Schaeper U., la Regina M., Bayley S.,
RA      Chinnadural G.;
RT      "A region in the C-terminus of adenovirus 2/5 E1a protein is required

```

```

RT      for association with a cellular phosphoprotein and important for the
RT      negative modulation of T24-ras mediated transformation, tumorigenesis
RT      and metastasis.";
RN      EMBL J. 12:469-478(1993).
CC      -1- FUNCTION: CO-REPRESSOR TARGETING DIVERSE TRANSCRIPTION REGULATORS.
CC      -1- SUBUNIT: INTERACTS WITH THE C-TERMINUS OF ADENOVIRUS E1A PROTEIN,
CC      E1K3 AND CTIP VIA THEIR CONSENSUS MOTIF P-X-(DNS)-L-(STVA).
CC      CAN FORM HOMODIMERS OR HETERODIMERS OF CTBP1 AND CTBP2.
CC      -1- SUBCELLULAR LOCATION: Nuclear (by similarity).
CC      -1- PTM: PHOSPHORYLATED. THE LEVEL OF PHOSPHORYLATION APPEARS TO BE
CC      REGULATED DURING THE CELL CYCLE.
CC      -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC      DEHYDROGENASES FAMILY, BUT HIGHLY DIVERGENT.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: U37408; AAC62822.1; -
DR      EMBL: AF091555; AAD14597.1; -
DR      EMBL: BC011655; AAH11655.1; -
DR      PDB: 1MX3; 18-DEC-02.
DR      Genew: HGNC:2494; CTBP1.
DR      MIM: 602618; -
DR      GO: GO:0008022; F:protein C-terminus binding activity; TAS.
DR      GO: GO:0008181; F:tumor suppressor; TAS.
DR      GO: GO:0008285; P:negative regulation of cell proliferation; TAS.
DR      GO: GO:0007048; P:oncogenesis; TAS.
DR      GO: GO:0006468; P:protein amino acid phosphorylation; TAS.
DR      GO: GO:0008166; P:viral replication; TAS.
DR      InterPro: IPR006139; 2-Hacid_DH.
DR      InterPro: IPR006140; 2-Hacid_DH_C.
DR      Pfam: PF00389; 2-Hacid_DH; 1.
DR      Pfam: PF02826; 2-Hacid_DH_C; 1.
DR      PROSITE: PS00065; D_2-HYDROXYACID_DH_1; 1.
DR      PROSITE: PS00670; D_2-HYDROXYACID_DH_2; FALSE_NEG.
DR      PROSITE: PS00671; D_2-HYDROXYACID_DH_3; FALSE_NEG.
DR      Nuclear protein; Phosphorylation; 3D-structure.
SQ      SEQUENCE 440 AA; 47535 MW; F071DD308385603F CRC64;

Query Match      14.4%; Score 307.5; DB 1; Length 440;
Best Local Similarity 30.8%; Pred. No. 2,1e-18;
Matches 94; Conservative 51; Mismatches 129; Indels 31; Gaps 6;

Qy      58 LGLRPLESNG-RTLVVTSKDPDSVFE-----RELVDADV 94
Db      13 LGVRPPI-MNGSLPRPLRVALLDGRCVEMPIIKDVAATVAFCDNOSTOETHEVYLEAV 71
Qy      95 ISQFVPAVYLPERIAKKNLKLATAGISGDHYDLSAIDRNTVVAEFTYCNISVAEH 154
Db      72 GALVYHTTLTREDLEKFKALRIIVRIGSGFDNDIKSAGDLGIAVCVPAASVEFADS 131
Qy      155 VVMATLSLVNRYLDSHEKARKG-----WNTADCVSNATYDEAMNHYGTVAAGRIGLAVLR 209
Db      132 TLCHILNRYRATWLHQALRGSTRVQSVQEQREVAAGARIRGEGTLGITGGRVGOAVL 191
Qy      210 RLAFVDVHLNHTDRRLPESVEKELENTLWNTATREDMYRVCVGNVLNCPILPETHMIND 269
Db      192 RAKTFGFEVNFYEDPY-LSDGIERALGLQRIYSTLDLFLHSCVTLHGLGNHNHNLNDE 250
Qy      270 TLKLFKRGATVNTARCKLDRDVAVALRESGLAGVDFPOP-APKDHPRTPMPYN 328
Db      251 TIKOMRGAGFLVNTARGSLVDKALQALKEGRIGRGAALDVHESEPSFGQPLKDAVNL 310
Qy      329 GMTPH 333
Db      311 ICTPH 315

```

```

RESULT 15
CTPI_MOUSE STANDARD; PRT; 440 AA.
ID CTPI_MOUSE
AC 088712; 090K2;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE C-terminal binding protein 1 (Ctbp1).
GN CTBP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=99298191; PubMed=10369679;
RX Crigui-Fillipe P., Ducret C., Maiza S.-M., Wasylyk B.;
RT "Net, a negative Ras-switchable TCF, contains a second inhibition
RT domain, the CID, that mediates repression through interactions with
RT CTBP and de-acetylation."
RL EMBO J. 18:3392-3403(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR;
RX MEDLINE=20036814; PubMed=10567582;
RA Furusawa T., Moribe H., Kondoh H., Higashi Y.;
RT "Identification of Ctbp1 and Ctbp2 as corepressors of zinc finger-
RT homeodomain factor deltaepl."
RL Mol. Cell. Biol. 19:8581-8590(1999).
CC -1- FUNCTION: CO-REPRESSOR TARGETING DIVERSE TRANSCRIPTION REGULATORS.
CC -1- SUBUNIT: INTERACTS WITH THE C-TERMINUS OF ADENOVIRUS E1A PROTEIN,
CC ELK3 AND CTIP VIA THEIR CONSENSUS MOTIF P-X-[DNS]-L-[STVA].
CC CAN FORM HOMODIMERS OR HETERODIMERS OF CTBP1 AND CTBP2
CC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT THE DEVELOPMENTAL STAGES
CC AND IN A WIDE RANGE OF ADULT TISSUES.
CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC DEHYDROGENASES FAMILY. BUT HIGHLY DIVERGENT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: A1010483; CA09219.1; -
DR EMBL: AB033122; BA085180.1; -
DR HSSP: P17584; IDXY.
DR MGD: MGI:1201685; Ctbp1.
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH_C.
DR Pfam: PF00389; 2-Hacid_DH.1.
DR Pfam: PF02826; 2-Hacid_DH.C.1.
DR PROSITE: PS00065; D_2_HYDROXYACID_DH.1.1.
DR PROSITE: PS00670; D_2_HYDROXYACID_DH.2. FALSE_NEG.
DR PROSITE: PS00671; D_2_HYDROXYACID_DH.3. FALSE_NEG.
KM Nuclear protein.
FT CONFLICT 379 379 Y -> YS (TN REF. 2).
SQ SEQUENCE 440 AA; 47657 MW; C51D5894DC594617 CRC64;

Query Match 14.4%; Score 306.5; DB 1; Length 440;
Best Local Similarity 30.5%; Pred. No. 2,5e-18;
Matches 93; Conservative 52; Mismatches 129; Indels 31; Gaps 6;

QY 58 LGRLRYLESNG--HTLVVTSKDKDPDSYFE-----RELVDADV 94
DB 13 LGVRPPI-MNGRPMHPRPLVALLDGRDCTVEMPIKDVATVAFCDASTOIRHKKVLINEAV 71
QY 95 ISQPFMPAVYLTPEIRAKAKLKLATAGIGSDHVDLOSADIRNVTVAAEVTYCNSISVAEH 154
DB 95 ISQPFMPAVYLTPEIRAKAKLKLATAGIGSDHVDLOSADIRNVTVAAEVTYCNSISVAEH 154

```

```

DB 72 GALTMYHTITVLTREDLEKFKALRIIVRIGSGFDNIDIKSAGDLCIANCNPASVEETAOS 131
QY 155 VMAMILSLVBNVYLPSEHMARKG-----ANIADCVSHAYDLEAMHYGTVAAGRIGLAVLR 209
DB 132 TLCHILNLVYRTTWLHGLAREGTRVOSVEQIREVASAAMIRRETLGIIIGLGVCAVAL 191
QY 210 RLAPFDVHLHYTDRLRPESVEKELNTWHTATREDMYPVCDVYTLNCPLEPETEHMINDE 269
DB 192 RAKAFGFNVLFYDPY-LSDGIERALGIQVRSTLQDLFLHSDCVTLHCGLENNHHLINDF 250
QY 270 TLKLFKRGAYIVMTARGKLCDDRDVAVARALESGRAGYAGVWFPPOP-APKDHPRMPVYN 328
DB 251 TVKQMRGAFVMTARGGLVDERKALQALKEGRIRGALDVHSEPSFSQGPLKADPNL 310
QY 329 GMTPH 333
DB 311 ICTPH 315

```

Search completed: July 24, 2003, 20:46:32
 Job time : 26 secs